

## Search Notes

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

On nucleic - nucleic search, using SW model

Run on: March 1, 2005, 18:09:49 ; Search time 1816 Seconds  
(without alignments)  
480.283 Million cell updates/sec

Title: US-10-677-982-1

Perfect score: 18

Sequence: 1 GTTCTACATAATGCGCCG 18

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

database : GenBank:  
1: gb\_ba::\*  
2: gb\_htg::\*  
3: gb\_in::\*  
4: gb\_om::\*  
5: gb\_ov::\*  
6: gb\_pat::\*  
7: gb\_ph::\*  
8: gb\_pi::\*  
9: gb\_pr::\*  
10: gb\_O::\*  
11: gb\_stb::\*  
12: gb\_SY::\*  
13: gb\_un::\*  
14: gb\_vl::\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query | Match Length | DB | ID        | Description                                 | RESULT | ALIGMENTS                  |
|------------|-------|-------|--------------|----|-----------|---|--------|----------------------------|
| 1          | 18    | 100.0 | 18           | 6  | AR534202  | AR534202 Sequence 1 from patent US 6733971. | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 2          | 18    | 100.0 | 18           | 6  | AR534203  | AR534203 Sequence 1 from patent US 6733971. | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 3          | 18    | 100.0 | 173185       | 10 | AC125317  | AC125317 Mus musculus                       | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 4          | 18    | 100.0 | 190369       | 10 | AC119908  | AC119908 Mus musculus                       | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 5          | 16.4  | 91.1  | 677          | 8  | AK103255  | AK103255 Drosophila                         | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 6          | 16.4  | 91.1  | 47636        | 2  | AC014385  | AC014385 Drosophila                         | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 7          | 16.4  | 91.1  | 92958        | 2  | DMBR11J17 | DMBR11J17                                   | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 8          | 16.4  | 91.1  | 144233       | 8  | AP003240  | AP003240 Oryza sativa                       | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 9          | 16.4  | 91.1  | 165928       | 3  | AC023725  | AC023725 Drosophila                         | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 10         | 16.4  | 91.1  | 172784       | 3  | AC105352  | AC105352 Drosophila                         | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 11         | 16.4  | 91.1  | 181831       | 3  | AC107403  | AC107403 Drosophila                         | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 12         | 16.4  | 91.1  | 309037       | 3  | AB003429  | AB003429 Drosophila                         | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 13         | 16.4  | 91.1  | 321355       | 2  | AC128492  | AC128492 Rattus norvegicus                  | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 14         | 16.4  | 88.9  | 303750       | 1  | AB016931  | AB016931 Bacteroides                        | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 15         | 15.4  | 85.6  | 5459         | 8  | AK03186   | AK03186 Oryza sativa                        | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 16         | 15.4  | 85.6  | 855          | 8  | AK056807  | AK056807 Oryza sativa                       | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 17         | 15.4  | 85.6  | 1519         | 6  | CQ593598  | CQ593598 Sequence                           | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 18         | 15.4  | 85.6  | 2217         | 3  | AY052136  | AY052136 Drosophila                         | 18 bp  | DNA linear PAT 08-OCT-2004 |
| 19         | 15.4  | 85.6  | 2336         | 3  | BT001590  | BT001590 Drosophila                         | 18 bp  | DNA linear PAT 08-OCT-2004 |

**TITLE** Method for identifying an agent that affects a hedgehog signaling pathway

**JOURNAL** Patent: US 6733971-A 2 11-MAY-2004;

**FEATURES** Location/Qualifiers

**source** I..18

**ORIGIN** /organism="Unknown" /mol\_type="Genomic DNA"

**RESULT 3** Query Match

Best Local Similarity 100.0%; Score 18; DB 6; Length 18;

Matches 18; Consecutive 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTCATACATATGCGCC 18

Db 18 GTCTACATATGCGCC 1

**DEFINITION** Mus musculus BAC clone RP24-286L23 from chromosome 7, complete sequence.

**ACCESSION** ACI25317

**VERSION** ACI25317.4 GI:28475488

**KEYWORDS** HTG.

**SOURCE** Mus musculus (house mouse)

**ORGANISM** Mus musculus

**REFERENCE** Kruchowski, S., Meyer, R. and Doeberl, A.

**AUTHORS** 1 (bases 1 to 173185)

**JOURNAL** Mammalia; Butcheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE** Mammalia; Butcheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**AUTHORS** 2 (bases 1 to 173185)

**JOURNAL** Unpublished (2001)

**REFERENCE** McPherson, J.D. and Waterston, R.H.

**AUTHORS** 3 (bases 1 to 173185)

**JOURNAL** Direct Submission

**REFERENCE** Submitted (22-JUN-2002) Genome Sequencing Center, 4444 Forest Park

**AUTHORS** Parkway, St. Louis, MO 63108, USA

**JOURNAL** Unpublished (2001)

**REFERENCE** McPherson, J.D. and Waterston, R.H.

**AUTHORS** 4 (bases 1 to 173185)

**JOURNAL** Submitted (22-NOV-2002) Genome Sequencing Center, 4444 Forest Park

**REFERENCE** Parkway, St. Louis, MO 63108, USA

**AUTHORS** McPherson, J.D. and Waterston, R.H.

**JOURNAL** Direct Submission

**REFERENCE** Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park

**AUTHORS** 5 (bases 1 to 173185)

**JOURNAL** Direct Submission

**REFERENCE** Parkway, St. Louis, MO 63108, USA

**AUTHORS** 6 (bases 1 to 173185)

**JOURNAL** Direct Submission

**COMMENT** Submitted (13-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Feb 23, 2003 this sequence version replaced gi:25167251.

----- Genome Center

Center: Washington University Genome Sequencing Center

Web site: <http://genome.wustl.edu>

Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)

----- Center project name: M\_BB0286L23

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

**TITLE** Method for identifying an agent that affects a hedgehog signaling pathway

**JOURNAL** Patent: US 6733971-A 2 11-MAY-2004;

**FEATURES** Location/Qualifiers

**source** I..18

**ORIGIN** /organism="Unknown" /mol\_type="Genomic DNA"

**RESULT 3** Query Match

Best Local Similarity 100.0%; Score 18; DB 6; Length 18;

Matches 18; Consecutive 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTCATACATATGCGCC 18

Db 18 GTCTACATATGCGCC 1

**DEFINITION** Mus musculus BAC clone RP24-286L23 from chromosome 7, complete sequence.

**ACCESSION** ACI25317

**VERSION** ACI25317.4 GI:28475488

**KEYWORDS** HTG.

**SOURCE** Mus musculus (house mouse)

**ORGANISM** Mus musculus

**REFERENCE** Kruchowski, S., Meyer, R. and Doeberl, A.

**AUTHORS** 1 (bases 1 to 173185)

**JOURNAL** Mammalia; Butcheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE** Mammalia; Butcheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**AUTHORS** 2 (bases 1 to 173185)

**JOURNAL** Unpublished (2001)

**REFERENCE** McPherson, J.D. and Waterston, R.H.

**AUTHORS** 3 (bases 1 to 173185)

**JOURNAL** Direct Submission

**REFERENCE** Submitted (22-JUN-2002) Genome Sequencing Center, 4444 Forest Park

**AUTHORS** Parkway, St. Louis, MO 63108, USA

**JOURNAL** Unpublished (2001)

**REFERENCE** McPherson, J.D. and Waterston, R.H.

**AUTHORS** 4 (bases 1 to 173185)

**JOURNAL** Submitted (22-NOV-2002) Genome Sequencing Center, 4444 Forest Park

**REFERENCE** Parkway, St. Louis, MO 63108, USA

**AUTHORS** McPherson, J.D. and Waterston, R.H.

**JOURNAL** Direct Submission

**REFERENCE** Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park

**AUTHORS** 5 (bases 1 to 173185)

**JOURNAL** Direct Submission

**REFERENCE** Parkway, St. Louis, MO 63108, USA

**AUTHORS** 6 (bases 1 to 173185)

**JOURNAL** Direct Submission

**COMMENT** Submitted (13-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Feb 23, 2003 this sequence version replaced gi:25167251.

----- Genome Center

Center: Washington University Genome Sequencing Center

Web site: <http://genome.wustl.edu>

Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)

----- Center project name: M\_BB0286L23

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

**MAPPING INFORMATION:** Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

**SOURCE INFORMATION:** The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

**FEATURES** This sequence is the entire insert of the clone.

**source** Location/Qualifiers 1..173185

**ORGANISM** Mus musculus

**mol\_type** "Genomic DNA"

**db\_xref** "taxon:10090"

**chromosome** "7"

**map** "7"

**clone** "RP24-286L23"

**clone\_lib** "RPCI-24"

**..1145**

**rpt\_family** "L1"

**1413..1655**

**rpt\_family** "M4LR"

**1901..3478**

**rpt\_family** "L1"

**4384..4436**

**rpt\_family** "SRV1"

**4465..4750**

**rpt\_family** "B4"

**4751..5143**

**rpt\_family** "M4LR"

**5144..5168**

**rpt\_family** "B4"

**5335..5627**

**rpt\_family** "B4"

**6055..6117**

**rpt\_family** "ID"

**6792..6880**

**rpt\_family** "Alu"

**7466..7854**

**rpt\_family** "L2"

**7943..8168**

**rpt\_family** "B4"

**8170..8400**

**rpt\_family** "L1"

**8404..8680**

**rpt\_family** "L1"

**9031..9202**

**rpt\_family** "B4"

**10342..10527**

**rpt\_family** "SRV1"

**10630..10704**

**rpt\_family** "ID"

**10736..10855**

**rpt\_family** "Alu"

**11059..11143**

**rpt\_family** "Alu"

**1137..11520**

**rpt\_family** "B2"

**11805..11940**

**rpt\_family** "M4LR"

repeat\_region 12555 .12729 /rpt:family="B2"  
repeat\_region 12794 .13058 /rpt:family="MER1\_type"  
repeat\_region 13551 .13759 /rpt:family="B4"  
repeat\_region 13747 .14155 /rpt:family="L1"  
repeat\_region 14169 .14359 /rpt:family="B2"  
repeat\_region 14395 .14711 /rpt:family="B2"  
repeat\_region 14753 .14855 /rpt:family="ERVK"  
repeat\_region 14836 .15263 /rpt:family="L1"  
repeat\_region 15313 .15392 /rpt:family="ID"  
repeat\_region 15404 .15594 /rpt:family="L1"  
repeat\_region 17002 .17241 /rpt:family="L1"  
repeat\_region 17309 .17704 /rpt:family="L1"  
repeat\_region 17705 .17855 /rpt:family="L1"  
repeat\_region 18241 .18487 /rpt:family="B4"  
repeat\_region 18427 .18499 /rpt:family="B2"  
repeat\_region 18549 .18814 /rpt:family="L1"  
repeat\_region 18872 .18983 /rpt:family="L1"  
repeat\_region 19014 .19501 /rpt:family="L1"  
repeat\_region 21526 .21724 /rpt:family="B2"  
repeat\_region 22573 .22982 /rpt:family="MaLR"  
repeat\_region 22290 .23587 /rpt:family="L1"  
repeat\_region 25375 .25517 /rpt:family="B2"  
repeat\_region 26791 .26994 /rpt:family="B2"  
repeat\_region 27298 .27399 /rpt:family="Alu"  
repeat\_region 27635 .27811 /rpt:family="B2"  
repeat\_region 28131 .28312 /rpt:family="MaLR"  
repeat\_region 28378 .28483 /rpt:family="MaLR"  
repeat\_region 28910 .29053 /rpt:family="B4"  
repeat\_region 31816 .31952 /rpt:family="Alu"  
repeat\_region 32142 .32250 /rpt:family="L1"  
repeat\_region 32685 .32994 /rpt:family="MaLR"  
repeat\_region 33298 .33370 /rpt:family="L1"  
repeat\_region 33371 .34354 /rpt:family="L1"  
repeat\_region 34355 .34840 /rpt:family="L1"  
repeat\_region 34976 .35171 /rpt:family="B2"  
repeat\_region 36213 .37443 /rpt:family="L1"  
repeat\_region 37526 .37669 /rpt:family="Alu"

Query Match 100.0%; Score 18; DB 10; Length 173185;  
Best Local Similarity 100.0%; Pred. No. 24; Matches 18;  
Matched Conservative 0; Mismatches 0; Indels 0; GapB 0;  
QY 1 GTCTCTACATAATGCGCC 18  
DB 163463 GTCTCTACATAATGCGCC 163446

RESULT 4 AC119908  
LOCUS AC119908 190369 bp DNA linear ROD 08-SEP-2004  
DEFINITION Mus musculus chromosome 7, clone RP24-240G4, complete sequence.  
ACCESSION AC119908  
VERSION AC119908.16 GI:51921399  
KEYWORDS HTG.

SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
AUTHORS Birren,B., Nusbaum,C. and Lander,E.  
REFERENCE I (bases 1 to 190369)  
JOURNAL Unpublished  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukagaltz,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearallano,K., Dewart,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHughon,W., Gadea,D., Galagan,J., Godyn,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Keilis,C., Larocque,K., Lamazares,R., Landers,T., Lebocky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., McDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McElwain,P., McKernan,K., Medrini,J., Meneus,L., Mihowa,T., Mlenga,V., Murphy,T., Mayor,J., Nguyen,C., Nicoll,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Resta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vasilev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission  
JOURNAL Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
REFERENCE J (bases 1 to 190369)  
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,I., Boguslavsky,L., Boukagaltz,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearallano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyne,S., Graham,L., Grand-Pierre,N., Hafetz,N., Hagojan,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Keilis,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., McDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihowa,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Rettig,R., Riese,P., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vasilev,H., Vankataraman,V.S., Vie,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,





| QY                                      | 1  | GTTCATACATATGGCCG                                      | 18                         | 7503: gap of unknown length   |
|---|--|--|----------------------------|-------------------------------|
| Db                                      | 14262  | GTTCATACATATGGCCG                                      | 14245                      | 7504: contig of 218 in length |
| DEFINITION                              | DMBR11J17  | Drosophila melanogaster                                | 92958 bp                   | DNA linear HTG 11-OCT-1999    |
| ACCESSION                               | AL121811   | chromosome X clone BAPR11J17                           | map 4A strain              | 7721: contig of 90 in length  |
| VERSION                                 | AL121811.1   | SEQUENCING IN PROGRESS                                 | ***, 178 unordered pieces. | 7722: gap of unknown length   |
| KEYWORDS                                | RTG: HTGS_PHSEx1   |  |                            | 7821: gap of unknown length   |
| SOURCE                                  | Drosophila   | Drosophila melanogaster (fruit fly)                    |                            | 7911: contig of 90 in length  |
| ORGANISM                                | Drosophila melanogaster  |  |                            | 7912: gap of unknown length   |
| REMARK                                  | Bularyok, Neptera; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidae; Drosophilidae; Drosophila; RA Murphy, L., Harris, D. and Barrell, B.   |  |                            | 8011: gap of unknown length   |
| REFERENCE                               | 1 (bases 1 to 92958)   | Sequencing the X chromosome of Drosophila melanogaster |                            | 8012: gap of unknown length   |
| AUTHORS                                 | RA Murphy, L., Harris, D. and Barrell, B.  |  |                            | 8472: gap of unknown length   |
| TITLE                                   | Unpublished  |  |                            | 8473: gap of unknown length   |
| JOURNAL                                 | Submitted (10-OCT-1999) European Drosophila Genome Sequencing Consortium   |  |                            | 8576: contig of 164 in length |
| COMMENT                                 | This is a 'working draft' sequence. It currently consists of 178 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence. 1 415: contig of 415 in length |  |                            | 8737: gap of unknown length   |
| 1 416: contig of 415 in length          | 5 15: gap of unknown length  |  |                            | 8837: gap of unknown length   |
| 5 16: 953: contig of 438 in length      | 10 515: gap of unknown length  |  |                            | 9144: contig of 308 in length |
| 10 54: 1308: contig of 255 in length    | 14 1309: gap of unknown length   |  |                            | 9145: gap of unknown length   |
| 14 1408: gap of unknown length          | 18 1409: contig of 40 in length  |  |                            | 9146: gap of unknown length   |
| 18 1448: contig of 352 in length        | 22 1449: gap of unknown length   |  |                            | 9147: gap of unknown length   |
| 22 1549: 2026: contig of 478 in length  | 26 1549: gap of unknown length   |  |                            | 9148: gap of unknown length   |
| 26 2126: gap of unknown length          | 30 2126: contig of 200 in length   |  |                            | 9149: gap of unknown length   |
| 30 2127: 2326: contig of 32 in length   | 34 2127: gap of unknown length   |  |                            | 9150: gap of unknown length   |
| 34 2327: 2778: contig of 352 in length  | 38 2327: gap of unknown length   |  |                            | 9151: gap of unknown length   |
| 38 2779: 3125: contig of 247 in length  | 42 2779: gap of unknown length   |  |                            | 9152: gap of unknown length   |
| 42 3126: 3225: gap of unknown length    | 46 3126: contig of 32 in length  |  |                            | 9153: contig of 389 in length |
| 46 3226: 3357: contig of 32 in length   | 50 3226: gap of unknown length   |  |                            | 9154: contig of 341 in length |
| 50 3358: 3720: contig of 363 in length  | 54 3358: gap of unknown length   |  |                            | 9155: contig of 855 in length |
| 54 3721: 3820: gap of unknown length    | 58 3721: contig of 205 in length   |  |                            | 9156: contig of 341 in length |
| 58 3821: 4025: contig of 432 in length  | 62 3821: gap of unknown length   |  |                            | 9157: contig of 345 in length |
| 62 4026: 4125: gap of unknown length    | 66 4026: contig of 450 in length   |  |                            | 9158: contig of 492 in length |
| 66 4126: 4575: contig of 450 in length  | 70 4126: gap of unknown length   |  |                            | 9159: contig of 492 in length |
| 70 4576: 4676: gap of unknown length    | 74 4576: contig of 335 in length   |  |                            | 9160: contig of 492 in length |
| 74 4676: 4972: contig of 297 in length  | 78 4676: gap of unknown length   |  |                            | 9161: contig of 492 in length |
| 78 4973: 5072: gap of unknown length    | 82 4973: contig of 432 in length   |  |                            | 9162: contig of 492 in length |
| 82 5073: 5504: contig of 400 in length  | 86 5073: gap of unknown length   |  |                            | 9163: contig of 492 in length |
| 86 5505: 5664: gap of unknown length    | 90 5505: contig of 362 in length   |  |                            | 9164: contig of 492 in length |
| 90 5665: contig of 362 in length        | 94 5665: gap of unknown length   |  |                            | 9165: contig of 492 in length |
| 94 5666: 6066: gap of unknown length    | 98 5666: contig of 335 in length   |  |                            | 9166: contig of 492 in length |
| 98 6067: 6401: contig of 335 in length  | 102 6067: gap of unknown length  |  |                            | 9167: contig of 492 in length |
| 102 6402: 6501: gap of unknown length   | 106 6402: contig of 297 in length  |  |                            | 9168: contig of 492 in length |
| 106 6502: 6901: contig of 400 in length | 110 6502: gap of unknown length  |  |                            | 9169: contig of 492 in length |
| 110 6902: 7001: gap of unknown length   | 114 6902: contig of 23 in length   |  |                            | 9170: contig of 492 in length |
| 114 7002: 7024: contig of 23 in length  | 118 7024: gap of unknown length  |  |                            | 9171: contig of 492 in length |
| 118 7025: 7124: gap of unknown length   | 122 7025: contig of 279 in length  |  |                            | 9172: contig of 492 in length |
| 122 7124: 7403: contig of 279 in length |  |  |                            | 9173: contig of 492 in length |

|                         |  |   |
|-------------------------|--|---|
| 22027                   | 22323: contig of 297 in length                     | RESULT 8  |
| 22324                   | 22423: gap of unknown length                       | AP003240  |
| 22424                   | 22884: contig of 461 in length                     | AP003240  |
| 22885                   | 22984: gap of unknown length                       | Oryza sativa (Japonica cultivar-group) genomic DNA, chromosome 1, PAC clone: P0406G08.  |
| 22985                   | 23168: contig of 184 in length                     | AP003240  |
| 23169                   | 23268: gap of unknown length                       | AP003240.2  |
| 23269                   | 23578: contig of 310 in length                     | GI:1702605  |
| 23579                   | 23678: gap of unknown length                       |   |
| 23988                   | 23987: contig of 309 in length                     |   |
| 24087                   | 24087: gap of unknown length                       |   |
| 24427                   | 24427: contig of 340 in length                     |   |
| 24428                   | 24527: gap of unknown length                       |   |
| 24528                   | 25057: contig of 530 in length                     |   |
| 25058                   | 25157: gap of unknown length                       |   |
| 25158                   | 25777: contig of 620 in length                     |   |
| 25778                   | 25877: gap of unknown length                       |   |
| 25878                   | 26430: contig of 553 in length                     |   |
| 26431                   | 26531: gap of unknown length                       |   |
| 26531                   | 26965: contig of 435 in length                     |   |
| 26966                   | 27065: gap of unknown length                       |   |
| 27066                   | 27853: contig of 788 in length                     |   |
| 27854                   | 27953: gap of unknown length                       |   |
| 27954                   | 28287: contig of 334 in length                     |   |
| 28288                   | 28388: gap of unknown length                       |   |
| 28388                   | 28852: contig of 465 in length                     |   |
| 28853                   | 28952: gap of unknown length                       |   |
| 28953                   | 29147: contig of 195 in length                     |   |
| 29148                   | 29247: gap of unknown length                       |   |
| 29248                   | 29585: contig of 338 in length                     |   |
| 29586                   | 29885: gap of unknown length                       |   |
| 29886                   | 29830: contig of 145 in length                     |   |
| 29831                   | 29930: gap of unknown length                       |   |
| 29931                   | 30137: contig of 207 in length                     |   |
| 30138                   | 30237: gap of unknown length                       |   |
| 30238                   | 30524: contig of 287 in length                     |   |
| 30525                   | 30624: gap of unknown length                       |   |
| 30625                   | 30933: contig of 309 in length                     |   |
| 30934                   | 31033: gap of unknown length                       |   |
| 31034                   | 31458: contig of 425 in length                     |   |
| 31459                   | 31558: gap of unknown length                       |   |
| 31559                   | 31902: contig of 344 in length                     |   |
| 31903                   | 32002: gap of unknown length                       |   |
| 32003                   | 32778: contig of 776 in length                     |   |
| 32779                   | 32878: gap of unknown length                       |   |
| 32879                   | 33449: contig of 571 in length                     |   |
| 33450                   | 33549: gap of unknown length                       |   |
| 33550                   | 34007: contig of 458 in length                     |   |
| 34008                   | 34107: gap of unknown length                       |   |
| 34108                   | 34871: contig of 764 in length                     |   |
| 34872                   | 34971: gap of unknown length                       |   |
| 34972                   | 35441: contig of 470 in length                     |   |
| 35442                   | 35541: gap of unknown length                       |   |
| 35542                   | 36137: contig of 596 in length                     |   |
| 36138                   | 36237: gap of unknown length                       |   |
| 36238                   | 36507: contig of 270 in length                     |   |
| 36508                   | 36607: gap of unknown length                       |   |
| 36608                   | 36821: contig of 214 in length                     |   |
| 36822                   | 36921: gap of unknown length                       |   |
| 36922                   | 37545: contig of 624 in length                     |   |
| 37546                   | 37645: gap of unknown length                       |   |
| 37646                   | 38331: contig of 686 in length                     |   |
| 38332                   | 38431: gap of unknown length                       |   |
| 38432                   | 38868: contig of 437 in length                     |   |
| Query                   | Match  | 91.1%; Score 164; DB 2; Best Local Similarity 94.4%; Prcd. No. 2.2e-02; Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0; |
| Qy                      | 1  | GTTCTACATATATGGCCCG 18  |
| Db                      | 73481  | GTTCTACATATATGGCCCG 73464   |
| FEATURES                | source   |   |
| 1. 1..4233              | /organism="Oryza sativa (Japonica cultivar-group)" |   |
| /mol_type="genomic DNA" | /cultivar="Nipponbare"                             |   |
| /db_xref="Raxxon:39947" | /chromosome="1"                                    |   |
| /clone="P0406G08"       | join(3501..3543,4068..4149,4324..4408)             |   |
| gene                    |  |   |

CDS

```

/gene="P0406G08.1"
join(3301. .3543. 4068. .4149. 4324. .4408)
/gene="P0406G08.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB86441.1"
/db_xref="GI:19571014"
/translation="WAPGRPIDIIDENFTGRKAIEGGAEDIAEAQKSTPSGDRARR
PRVIVIAMSIVGAHOOGRGRARDNP"
join(220. .6288. 6389. .6454. 6573. .6711. 7009. .7150,
7496. .7525)
/gene="P0406G08.2"
/codon_start=1
/product="putative FIEQL protein"
/protein_id="BAB86442.1"
/db_xref="GI:19571015"
/translation="WAGGVLVEHLVDAKGITGNDTGEICKIDPQVYVWVQYRSBRS
SVARDQKGPSWNEVKQINSAATOQHKUFLRIMHDTSRSDPFLGEATINTDLI
SIAMEHGTWEMSSKHRVVLADTYIGEIRVSLTFTASAKMNTIVES"
complement(join(9601. .9724. 10890. .1115. 11301. .11727,
12266. .12560. 12604. .12753. 12831. .1297. 13044,
13164. .13543. 13785. .14009. 14468. .14596. 14678. .14755,
14829. .14942. 15085. .15243. 15237. .16488. 171615. .17715)
/gene="P0406G08.3"
complement(join(9601. .9724. 10890. .1115. 11301. .11727,
12266. .12560. 12604. .12753. 12831. .1297. 13044,
13164. .13443. 13785. .14009. 14468. .14596. 14678. .14755,
14829. .14942. 15084. .15243. 16237. .16488. 171615. .17715)
/gene="P0406G08.3"
/note="contains ESTs AU057528(S21533), AU057527(S21533)"
```

CDS

```

/codon_start=1
/product="ribonuclease II-like protein"
/protein_id="BAB86443.1"
/db_xref="GI:19571016"
/translation="WAPGESSPLTRPAPPHUHALPPLRURSMSVAGAHSRPS
VLEERSQRGRGRVAKIGQGTRPLGRAVWHRSLYGGCGRGROVHGVD
DGGKQWVKVTDQNLISLTKPQVYVPGIINWHSDEFTKQDQDPTVACAW
AQEVSLRKLKLAKEAEGEMETVQLTSKAKDLPKPSKDSWMEKUKKESLQ
AYAVDADCDDEQRPAQVTKAMGRSKSSAKLKLINIGYEPVHNULDKRDIRY
TEELLSAANKEBLVQDSDKVTKDSTKVOYALDDEBDEDDLSAANLDPKIV
WIRHADPSLWORSITDREAMIGTSILPFTATPMPERLAMMASLQOGRCGKSV
TVSTILOPDGSKVNTICVVISATIAMEPIKVIENPDMYIESTELLYMNLBEE
LRIQEAMSLRAWRRSQVSIDTAMPIERKVIENPDMYIESTELLYMNLBEE
EMMILCGEVAARGNSNIALDQYRGIHOSNIVAVSAFTHLPEGPANSFANTS
DFOPKPREGVIGGIVGIVVNTFIRKQVNLQYQVNPFLLTFNADGRCTKQI
KAFGRGSSPFSYQAGDIEGMETFASHMKVAVLRLRNLWFLYLRQPKGKVKAL
IUKTRIKLAVLIVTIVGIGOTAVVSTGKGDETSVUVEAHHRRDILSVTEEV"
complement(join(18872. .19012. 20426. .20452))
```

gene

```

/gene="P0406G08.4"
complement(join(18872. .19012. 20426. .20452))
/gene="P0406G08.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB86444.1"
/db_xref="GI:19571017"
/translation="WVAAGFCMOKVYVLSRSRGCKLTIANNGHQSQIDSVLMPLPKSL
IVQYQNKFSIS"
join(21355. .21369. 21742. .21913. 23339. .23412)
/gene="P0406G08.5"
join(21355. .21369. 21742. .21913. 23339. .23412)
/gene="hypothetical protein"
/codon_start=1
/protein_id="BAB86445.1"
/db_xref="GI:19571018"
/translation="MPKAKTRTWHQIQLQGVVAHASTTDVFFLNQSPSTCFPAPSRN
RNCCELMILPSSTRSDEAVTAVLVLVYRLMFTVRIIDE"
join(24879. .24973. 25429. .25500. 25565. .25790)
```

CDS

```

/gene="P0406G08.6"
join(24879. .24973. 25429. .25500. 25565. .25790)
/gene="P0406G08.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB86446.1"
/db_xref="GI:19571019"
/translation="MDXINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
PRGAVGVBIGEFGQNSISVAREBPGICAV"
/translation="HGTIVHVPFQMDRACVUMPTWPKGPAAPLRTITPSADVLPLRASVRPACGSVSGG
28926. .31182
/gene="P0406G08.7"
28926. .31382
/CDS
/gene="P0406G08.7"
/note="contains ESTs AU162304(E60027), AU030624(E60027)
similar to SCARCRW gene regulator"
/codon_start=1
/protein_id="BAB86447.1"
/db_xref="GI:19571020"
/translation="WDAQYHPTWVTPMLPSKSKRDLPLPITYQTAANGFATAEFDPL
LFLSDAVCGGGSDYLNIVSAQPTISANSTNGASPRDVSTASASAMQAQHDSE
FSDIVLGYNRMMLADEKDEBKEHYPNADDLAAEKPFLTLDOSDPSGGSSESP
SGGAQVDPDPEFVOLKEVRLNPSOSSESSNSVTESEGESELSSPLPGDVS
SLANEFVQVQQLQFRFQEASKFLDDEBKSVKVIDTQKLGQKQKQKQKQKQKQ
VKVKEITSEVSAHGRGKHFQDGLDAAEKGSKHSAQGIDKQDHLVDRDMDKVLCN
GETCSKGKVKLRLDQVAKSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
DSQSSAANCNSLSCNCASSDSSCSDGKQKQKQKQKQKQKQKQKQKQKQKQKQ
LCTDVLKQYQVLAECPFKKSKHYPANOTLAVAKQHUTVYQGCL
QKUSNRDGGPPLKURITDTPGPFPAERETEYRQLSDYQAFNVPFQOIASRF
EVAYMEDLHETDEDVLCNGEFCFKLMDSVAVSPRNALKTQKMPHVTGHWG
NSGSYMAPFPTFPTFPTFPTFPTFPTFPTFPTFPTFPTFPTFPTFPTFPTF
LKERMEPTYKQWVNQVRNPKQFQQLPQDNQMMKQARKEVKRYHQFIDEDNWLQG
WKRCLPFLSTWPKDNRSS"
complement(33662. .33862)
/gene="P0406G08.8"
complement(33662. .33862)
/gene="P0406G08.8"
/translation="MDSTFCTSVFADKEDWAVQFSDTHGDSYTANAKLSPSLRK
KMWFKRTRKTRIAVFKRTRP"
complement(join(34857. .35053. 35343. .35412))
/gene="P0406G08.9"
complement(join(34857. .35053. 35343. .35412))
/gene="P0406G08.9"
/note="hypothetical protein"
/codon_start=1
/db_xref="GI:19571021"
/translation="MDSTFCTSVFADKEDWAVQFSDTHGDSYTANAKLSPSLRK
KMWFKRTRKTRIAVFKRTRP"
complement(join(34857. .35053. 35343. .35412))
/gene="P0406G08.9"
complement(join(34857. .35053. 35343. .35412))
/gene="P0406G08.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB86449.1"
/db_xref="GI:19571022"
/translation="MAILAKUENLASNTRHFLRECGDQDUDVTHDYGKKNPENCSCSS
SSRKSPTKTPDDGGRASPSVYIACDCTETHVTKRPPSLD"
complement(join(36419. .36549. 36655. .37629. 38130. .38258,
38660. .38743. .38864. 38995. .39096))
Query Match Best Local Similarity 91.1%; Score 16.4; DB 8; Length 144233;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

gene

```

QY 1 GTTCATACATATGGCCCG 18
Db 122554 GTTCATACATATGGCCCG 122571
```

RESULTS

| AC033725   | AC023725   | LOCUS  | AC023725                         |
|------------|------------|--|----------------------------------|
| DEFINITION | DEFINITION | Drosophila melanogaster X BAC Rb9-9-110 (Roswell Park Cancer Institute Drosophila BAC Library) | linear DNA linear IN 24-JUL-2002 |
| ACCESSION  | ACCESSION  | AC023725   | complete sequence.               |
| VERSION    | VERSION    | AC023725.5   | GT:21953977                      |

| KEYWORDS  | HTG.  |
|-----------|---|
| SOURCE    | Drosophila melanogaster (fruit fly)   |
| ORGANISM  | Drosophila melanogaster   |
| REFERENCE | 1 (bases 1 to 165928)   |
| AUTHORS   | Muzzny,D., Scherer,S., Adams,M.D., Holt,R.A., Evans,C.A., Gocayne,J.D., Tabor,P., Williamson,A., Housi,F.H., Dugan-Rocha,S.D., Sodergren,E.S., Hodgson,A.H., Chen,R.C., Ayotte,M., Scott,G.S., Worley,K.W., Amanatides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Beeson,K.Y., Brown,M., Buhay,C., Bush,D.A., Center,A., Chen,G., Chen,Z., Clerc-Blankenburg,K., Davenport,L.B., Dietz,S.M., Ding,Y., Dodson,K., Douc,L.E., Draper,H., Emery-Cohen,A., Ferriera,S., Garg,N.D.S., Houck,J., Hostin,D., Howland,T.J., Hume,J., Ihngwan,C., Jalali,M., Kovar,C., Liu,W., Maretz,B., McIntosh,T.C., Morgan,M., Moy,M., Murphy,B., Nelson,K.A., Ndasaa,Y., Perez,L., Pittman,G.S., Puris,V., Scheeler,P., Shen,H., Strong,R., Tector,C., Wang,Q., Williams,S.M., Wheeler,D., Weinstock,G., Gibbs,R. and Venter,J.C.   |
| TITLE     | Direct Submission   |
| JOURNAL   | Unpublished   |
| REFERENCE | 2 (bases 1 to 165928)   |
| AUTHORS   | Worley,K.C., Adams,C., Ado-Oduola,B., Ali-Osman,F.R., Allen,C., Alabrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Blimke,K., Blankenburg,C., Brem,B., Bonin,D., Bouck,J., Bowie,S., Brieva,M., Brown,B., Brown,M., Bryant,N.P., Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavacos,S.R., Chacko,D., Chaves,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dedeich,D.A., Delaney,K.R., Delgado,O., Dern,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earthart,C., Edgar,D., Edwards,C.C., Elhai,C., Escote,M., Falls,T., Ferragut,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Guitart,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J., Hernandez,O., Hodges,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J.J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichatoge,O., Lieu,C., Liu,J., Liu,W., Loulsaged,H., Loraod,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maeshiro,M., Mapia,P., Martin,R., Martindale,A., Martinez,B., Massey,B., Mawhinney,E., McLeod,M.P., Meadow,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Morris,M., Neal,D., Newton,J., Newtonson,N., Nguyen,A., Nguven,N., Nguven,N., Nickerson,E., Nwokonkwo,S., Oguh,M., Okunnu,G., Otagunye,N., Ovied,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Relias,R., Rejubon,I., Roffe,M., Rui,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoahtari,N., Sibson,I., Sodergren,E., Sonnai,T., Spark,R.A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansy,J., Taylor,C., Taylor,T., Telrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Wellington,C., Watlington,S., Williams,G., Williamson,A., Wlaczek,R., Wooden,S., Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R. |
| TITLE     | Direct Submission   |
| JOURNAL   | Unpublished   |
| REFERENCE | 3 (bases 1 to 165928)   |
| AUTHORS   | Worley,K.C.   |
| TITLE     | Direct Submission   |
| JOURNAL   | Submitted (17-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  |
| REFERENCE | 4 (bases 1 to 165928)   |
| AUTHORS   | BCM-HGSC.   |
| TITLE     | Direct Submission   |
| JOURNAL   | Submitted (20-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  |
| REFERENCE | 5 (bases 1 to 165928)   |
| AUTHORS   | BCM-HGSC.   |
| TITLE     | Direct Submission   |
| JOURNAL   | Submitted (01-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  |
| REFERENCE | 6 (bases 1 to 165928)   |
| AUTHORS   | BCM-HGSC.   |
| TITLE     | Direct Submission   |
| JOURNAL   | Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  |
| REFERENCE | 7 (bases 1 to 165928)   |
| AUTHORS   | BCM-HGSC.   |
| TITLE     | Direct Submission   |
| JOURNAL   | Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  |
| COMMENT   | On Jul 24, 2002 this sequence version replaced gi:21306516. INFORMATION: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a> or email: <a href="mailto:gc-help@bcm.tmc.edu">gc-help@bcm.tmc.edu</a>   |
| FEATURES  | <p><b>CLONE LENGTH:</b> This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.</p> <p><b>ANNOTATION OF FEATURES:</b> STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dSTS, GDB, and local mapping efforts.</p> <p>Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.</p> <p>Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3421). Similarity (expect &lt; 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.</p> <p><b>SEQUENCING READ COVERAGE:</b> Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.</p> <p><b>QUALITY OF INDIVIDUAL BASES:</b> This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <a href="http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html">http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html</a>.</p> <p>Location/Qualifiers</p> <p>1. .165928</p> <p>/organism="Drosophila melanogaster"</p> <p>/mol_type="genomic DNA"</p> <p>/chromosome="X"</p> <p>/clone="RPG8-9110"</p> <p>40958 .40968</p> <p>/notes="Baylor,14T/Celera,15T"</p> <p>/functions="polymorphic site"</p> <p>10316</p> <p>/function="low quality"</p> <p>101364 .101369</p>   |

## ORIGIN

/ Function="low quality"

| RESULT     | 10   | Query            | Match                  | 91.1%    | Score    | 16.4 | DB      | 3  | length     | 165928 |
|------------|--|------------------|------------------------|----------|----------|------|---------|----|------------|--------|
| LOCUS      | AC105352   | Best             | Local Similarity       | 94.4%    | Pred.    | NO.  | 2.3e+02 | ;  | Mismatches | 0;     |
| DEFINITION | Drosophila melanogaster  | 17;              | Conservative           | 0;       | Indels   | 0;   | Gaps    | 0; |            |        |
| ACCESSION  | AC105352   | Institute        | Drosophila BAC Library | complete | Sequence |      |         |    |            |        |
| VERSION    | AC105352.2   | HTG.             | GR:21392422            |          |          |      |         |    |            |        |
| KEYWORDS   |  |                  |                        |          |          |      |         |    |            |        |
| SOURCE     | Drosophila melanogaster  | (fruit fly)      |                        |          |          |      |         |    |            |        |
| ORGANISM   | Drosophila melanogaster  |                  |                        |          |          |      |         |    |            |        |
| REFERENCE  |  |                  |                        |          |          |      |         |    |            |        |
| AUTHORS    | Muzny, D., Scheier, S., Adams, M.D., Holt, R.A., Evans, C.A., Dugan-Roth, S.D., Sodergren, E.S., Hodgen, A.H., Chen, R.C., Ayele, M., Scott, G.S., Worley, K.W., Amanatidis, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Beeson, Y., Brown, M., Buahy, C., Busam, D.A., Center, A., Chen, G., Chen, Z., Clerc-Blankenburg, K., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, C., Carter, M., Carvazo, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dedekich, D.A., Delaney, K.R., Delgado, O., Dean, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Dreger, H., Dugan-Rocha, S., Durbin, K.J., Earnhardt, C., Edgar, D., Edwards, C.C., Elmarj, C., Escott, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.R., Foster, P., Franz, P., Gabisi, A., Gao, J., Garcia, J., Garner, T., Garsa, N., Gill, R., Gorrell, J.H., Guevara, W., Gunarane, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, C., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homai, F., Howard, G., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Kariisson, E., Kelly, S., Khan, U., King, L., Korvaj, J., Kovar, C., Kratovic, J., Kurese, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharse, O., Lieu, L., Liu, J., Liu, W., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, B., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, M., Moser, M., Neal, D., Newison, J., Newtonson, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, J., Newtonson, N., Oguh, M., Okwunou, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rajas, A., Rojubokan, I., Rolfe, M., |                  |                        |          |          |      |         |    |            |        |
| JOURNAL    | Unpublished  |                  |                        |          |          |      |         |    |            |        |
| REFERENCE  | 2 (bases 1 to 172784)  |                  |                        |          |          |      |         |    |            |        |
| AUTHORS    | Alsbrooks, S.L., Amaral, H.C., Are, J.R., Banks, T., Barbaria, J., Bowe, S., Brileva, M., Brown, B., Brown, M., Bryant, N.P., Buahy, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, C., Carter, M., Carvazo, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dedekich, D.A., Delaney, K.R., Delgado, O., Dean, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Dreger, H., Dugan-Rocha, S., Durbin, K.J., Earnhardt, C., Edgar, D., Edwards, C.C., Elmarj, C., Escott, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.R., Foster, P., Franz, P., Gabisi, A., Gao, J., Garcia, J., Garner, T., Garsa, N., Gill, R., Gorrell, J.H., Guevara, W., Gunarane, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, C., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homai, F., Howard, G., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Kariisson, E., Kelly, S., Khan, U., King, L., Korvaj, J., Kovar, C., Kratovic, J., Kurese, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharse, O., Lieu, L., Liu, J., Liu, W., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, B., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Morgan, M., Morris, M., Moser, M., Neal, D., Newison, J., Newtonson, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, J., Newtonson, N., Oguh, M., Okwunou, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rajas, A., Rojubokan, I., Rolfe, M.,   |                  |                        |          |          |      |         |    |            |        |
| QY         | 1  | GTTCATCATATGCGCG | 18                     |          |          |      |         |    |            |        |
| Db         | 134784   | GTTCATCATATGCGCG | 134801                 |          |          |      |         |    |            |        |

CLONE LENGTH: This sequence does not necessarily represent the full length of the clone. For more information, see the sequence version report at <http://www.hgsc.bcm.tmc.edu> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu).

sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

## ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts. Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintain continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8080/quality.info/genbank.annotation.html>.

## FEATURES

## FEATURES

source  
JOURNAL  
Unpublished

2 (bases 1 to 181831)  
Worley, K.C., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Scheeler, F., Shen, H., Strong, R., Tector, C., Wang, Q., William, S.M., Xiang, J., Zaveri, J.S., Zhou, J., Zorrilla, S., Smith, H.O., Wheeler, D., Weinstock, G., Gibbs, R. and Venter, J.C.  
Bentley, J., Bimonte, K., Blankenburg, K., Bonin, D., Bouck, J., Bowie, S., Brieva, M., Brown, S., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carter, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy, Carroll, L., Deedrich, D.A., Delaney, K.R., Delgado, O., Dunn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Farnhart, C., Edgar, D., Edwards, C.C., Elhai, C., Escoto, M., Falls, T., Ferragudo, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Gersbach, C., Gille, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodges, R., Hogues, M., Holloway, C., Hollings, B., Hombi, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, B., Kelly, S., Khan, U., King, L., Korvai, J., Kovari, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewin, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisiged, H., Lozada, R.J., Lux, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Makeshwar, M., Mapua, R., Martin, R., Martindale, A., Martinez, E., Massey, B., Mawhinney, E., McLeod, M.P., Meador, M., Meil, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newison, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokonkwo, S., Oguh, N., Okwuonu, G., Oragunyene, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, R., Pu, L.H., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Severy, S., Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E., Sonnake, T., Sparks, H., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williamson, G., Williamson, A., Wleczek, R., Wooden, S., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

ORIGIN  
JOURNAL  
Unpublished

Query Match 91.1%; Score 16,4; DB 3; Length 172784;  
Best Local Similarity 94.4%; Pred. No. 2,3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCATACATAATGGCCG 18  
/DB XREF="taxon:7227"  
(clone="RP9-45017"

RESULT 11  
AC107403/C  
LOCUS AC107403 181831 bp DNA linear INV 30-MAY-2002  
DEFINITION Drosophila melanogaster X BAC RP9-21M21 (Roswell Park Cancer  
Institute Drosophila BAC Library) complete sequence.  
ACCESSION AC107403  
VERSION AC107403.6  
KEYWORDS GI:20986392  
SOURCE RTG  
ORGANISM Drosophila melanogaster (fruit fly)  
Drosophila melanogaster (fruit fly)  
Drosophila; Metacoxa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydcoidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 181831)  
Muzny, D., Scherer, S., Adams, M.D., Holt, R.A., Evans, C.A.,  
Gocayne, J.D., Tabor, P., Williamson, A., Homsy, P.H., Hodson, A.H., Chen, R.C.,  
Dugan-Rocha, S.D., Sodergren, E.S., Hodson, A.H., Chen, R.C.,  
Ayotte, M., Scott, G.S., Worley, K.W., Anamides, P.G., Brandon, R.C.,  
Rogers, Y., An, H., Baldwin, D., Benson, K.Y., Brown, M., Buhay, C.,  
Bubam, D.A., Center, A., Chen, G., Chen, Z., Clerc-Blankenburg, K.,  
Davenport, L.B., Dietz, S.M., Ding, Y., Dodson, K., Douc, L.B.,  
Draper, H., Emery-Cohen, A., Ferreira, S., Garg, N.D.S., Houck, J.,  
Hotting, D., Howland, T.J., Hume, J., Ibegwam, C., Jaiali, M., Kovar, C.,  
Liu, W., Mattei, B., McIntosh, T.C., Morgan, M., Moy, M., Murphy, B.,

REFERENCE  
JOURNAL  
Unpublished

COMMENT  
On May 20, 2002 this sequence version replaced gi:20514404.





Consensus quality: 218055 bases at least Q30  
 Consensus Quality: 220048 bases at least Q20  
 Estimated insert size: 20101; sum-of-contigs estimation  
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation  
 \* NOTE: Estimated insert size may differ from sequence length.  
 \* (See [http://www.hgsc.bcm.tmc.edu/docs/Genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank/draft_data.html))  
 \* NOTE: This is a "working draft" sequence. It currently  
 \* consists of 3 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1. 319026: contig of 319026 bp in length  
 \* 319027 319126: gap of unknown length  
 \* 319127 320168: contig of 1042 bp in length  
 \* 320169 320268: gap of unknown length  
 \* 320269 32335: contig of 1087 bp in length.  
 FEATURES source  
 1. 321355  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /clone="CH230-166M22"  
 1. .1190  
 /note="wgs end extension  
 clone end:Sp6"  
 54939-.55369  
 /note="clone boundary  
 clone end:Sp6  
 site:EcorI  
 end sequence:BH301347"  
 complement(279655..279925)  
 /note="wgs end boundary  
 clone end:T7  
 site:EcorI  
 end sequence:BH301346"  
 282280..28738  
 /note="wgs end extension  
 317757..319026  
 /note="wgs end extension  
 clone\_end:T7;  
 misc\_feature  
 ORIGIN  
 Query Match 91.1%; Score 16.4; DB 2; Length 321355;  
 Best Local Similarity 94.4%; Pred. No. 2.4e-02; Mismatches 1;  
 Matches 17; Conservative 94%; Indels 0; Gaps 0;  
 QY 1 GTCCTACTAATGGCGCG 18  
 Db 248822 GTCCTACTAATGGCGCG 248805  
 RESULT 14  
 ACCESSION AE016931  
 LOCUS AE016931  
 DEFINITION Bacteroides thetaiotaomicron VPI-5482  
 VERSION AE016931 AE015928  
 KEYWORDS AE016931.1 GI:29338516  
 SOURCE  
 ORGANISM Bacteroides thetaiotaomicron VPI-5482  
 Bacteria; Bacteroides; Bacteroides (Class); Bacteroidales;  
 Bacteroididae; Bacteroides.  
 REFERENCE  
 AUTHOR(S) Xu,J., Bjursell,M.K., Hämrod,J., Deng,S., Carmichael,L.K.,  
 Chiang,H.C., Hooper,L.V. and Gordon,J.I.  
 TITLE A Genomic View of the Human Bacteroides thetaiotaomicron VPI-5482  
 CENTER Project name: GYHF  
 Center clone name: CH230-166M22  
 ----- Summary Statistics  
 Assembly program: phrap; version 0.990329  
 Consensus quality: 213847 bases at least Q40





/db\_xref="axon:39947"  
/clone="001-112-B04"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 549;  
Best Local Similarity 94.1%; Pred. No. 6 2e+02; Mismatches 1;  
Matches 16; Conservative 0; Indels 0; Gaps 0;  
Qy 2 TCTCTACATATGCCG 18  
Db 311 TCTCTACATATGCCG 327

Search completed: March 2, 2005, 03:35:49  
Job time : 1826 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

On nucleic - nucleic search, using SW model  
Run on: March 1, 2005, 17:32:45 ; Search time 426 Seconds  
(without alignments)  
250.130 Million cell updates/sec

Title: US-10-677-982-1

Perfect score: 18

Sequence: 1 GTTCTAACATAATGCGCCG 18

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 295987067 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseq1980s:\*

2: geneseq1990s:\*

3: geneseq2000s:\*

4: geneseq2001s:\*

5: geneseq2001bs:\*

6: geneseq2002bs:\*

7: geneseq2002bs:\*

8: geneseq2003bs:\*

9: geneseq2003bs:\*

10: geneseq2003cs:\*

11: geneseq2003cs:\*

12: geneseq2004bs:\*

13: geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No. Score Query Match Length DB ID Description

1 18 100.0 18 2 AAX86621

Aax86621 Sonic hedgehog response element nucleotide sequence.

2 15.4 85.6 1519 4 ABL16077

Abi16077 Drosophil

Abi16076 Drosophil

Continuation (2 of

Aax86622 Probe for

Aaa94156 Adenovirus

Aaa94133 Adenovirus

Ad60599 Corton CD

Abd15713 Pseudomon

Abd15682 Pseudomon

Abd15757 Pseudomon

Abd15782 Pseudomon

Abd850197 Bacterium

Abg855711 Bacterial

Abg85863 Human aut

Abg835825 Human aut

Abg835825 Human aut

Abg835874 Human aut

Abg835791 Human aut

Abg835757 Human aut

2 15.4 85.6 4927 4 ABL16076

Abi16076 Drosophil

Continuation (2 of

Aax86622 Probe for

Aaa94156 Adenovirus

Aaa94133 Adenovirus

Ad60599 Corton CD

Abd15713 Pseudomon

Abd15682 Pseudomon

Abd15757 Pseudomon

Abd15782 Pseudomon

Abd850197 Bacterium

Abg855711 Bacterial

Abg85863 Human aut

Abg835825 Human aut

Abg835874 Human aut

Abg835791 Human aut

Abg835757 Human aut

3 15.4 85.6 11000 3 AAB01489\_3

Aab01489\_3

Continuation (4 of

Continuation (2 of

4 15.4 85.6 11000 3 AAB01489\_4

Aab01489\_4

Continuation (5 of

5 15.4 85.6 11000 3 AAB01489\_5

Aab01489\_5

Continuation (5 of

6 15.4 85.6 11000 3 AAB01489\_6

Aab01489\_6

Continuation (5 of

7 15.4 85.6 11000 3 AAB01489\_7

Aab01489\_7

Continuation (5 of

8 15.4 85.6 11000 3 AAB01489\_8

Aab01489\_8

Continuation (5 of

9 15.4 85.6 11000 3 AAB01489\_9

Aab01489\_9

Continuation (5 of

10 15.4 85.6 11000 3 AAB01489\_10

Aab01489\_10

Continuation (5 of

11 15.4 85.6 11000 3 AAB01489\_11

CC transcription factor binds to a hedgehog response element such as the present sequence. Modulation of the phosphorylation of the transcription factor is used to control expression of target genes involved in hedgehog-mediated signaling pathway, specifically for regulating proliferation or differentiation of neuronal cells; treatment of proliferative disease (specifically basal cell carcinoma, medulloblastoma and meningioma) or bifida). Measuring the ratio of phosphorylated:dephosphorylated forms of transcription factor is used for diagnosis of hedgehog-mediated signaling pathway-mediated familial midline defects (specifically cyclopia or neural tube defects)

Sequence 18 BP; 4 A; 5 C; 4 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.6; Mismatches 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCTCTAGATAGCGCG 18  
Db 1 GGTCTACATATGCGCCG 18

RESULT 2  
ABLI6077/c

ID ABLI6077 standard; cDNA; 1519 BP.  
XX  
AC  
XX  
DT 26-MAR-2002 (first entry)

DB Drosophila melanogaster expressed polynucleotide SEQ ID NO 42713.  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
AC ABLI6076;  
XX  
DT 26-MAR-2002 (first entry)

DB Drosophila melanogaster expressed polynucleotide SEQ ID NO 42710.  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
OS Drosophila melanogaster.

XX  
PN WO200171042-A2.  
XX  
AC ABLI6076;  
XX  
DT 27-SEP-2001.

DB Drosophila melanogaster expressed polynucleotide SEQ ID NO 42711.  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
OS Drosophila melanogaster.

XX  
PN 23-MAR-2001; 2001WO-US009231.  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.

XX  
PA (PEKE ) PE CORP NY.  
XX  
PT Venter JC, Adams M, Li PW, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR P-PSDB; ABB7193.

XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX  
PS Claim 1: SEQ ID NO 42710; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLI6176, ABLJ3051), expressed DNA sequences (ABLU1840, ABLU16175) and the encoded proteins (ABB5773-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

XX  
SQ Sequence 4927 BP; 1288 A; 1106 C; 1081 G; 1452 T; 0 U; 0 Other;

XX  
CC The invention relates to an isolated nucleic acid detection reagent useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLU1840-ABLU16175), expressed DNA sequences (ABLU1840-ABLU16175) and the encoded proteins (ABB5773-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

XX  
Sequence 1519 BP; 306 A; 393 C; 410 G; 410 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 4; Length 1519;  
Best Local Similarity 94.1%; Pred. No. 1.5e+02; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTCTCATATAGCGCG 18  
Db 2620 TTCCACATATGCQCG 2604

RESULT 3

ID ABLI6076/c  
XX  
AC ABLI6076;  
XX  
DT 26-MAR-2002 (first entry)

DB Drosophila melanogaster expressed polynucleotide SEQ ID NO 42710.  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
OS Drosophila melanogaster.

XX  
PN WO200171042-A2.

XX  
AC ABLI6076;

XX  
DT 26-MAR-2002 (first entry)

DB Drosophila melanogaster expressed polynucleotide SEQ ID NO 42711.  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
OS Drosophila melanogaster.

XX  
PN 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.

XX  
PA (PEKE ) PE CORP NY.

XX  
PT Venter JC, Adams M, Li PW, Myers EW;

XX  
DR WPI; 2001-656860/75.

XX  
DR P-PSDB; ABB7193.

XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX  
PS Claim 1: SEQ ID NO 42710; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLU1840-ABLU16175), expressed DNA sequences (ABLU1840-ABLU16175) and the encoded proteins (ABB5773-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

XX  
SQ Sequence 4927 BP; 1288 A; 1106 C; 1081 G; 1452 T; 0 U; 0 Other;

XX  
CC The invention relates to an isolated nucleic acid detection reagent useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLU1840-ABLU16175), expressed DNA sequences (ABLU1840-ABLU16175) and the encoded proteins (ABB5773-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

XX  
Sequence 1519 BP; 306 A; 393 C; 410 G; 410 T; 0 U; 0 Other;

XX  
Sequence 1519 BP; 306 A; 393 C; 410 G; 410 T; 0 U; 0 Other;

WP ABA03041\_00 1 110000  
 WP ABA03041\_01 100001 210000  
 WP ABA03041\_02 200001 310000  
 WP ABA03041\_03 300001 410000  
 WP ABA03041\_04 400001 510000  
 WP ABA03041\_05 500001 610000  
 WP ABA03041\_06 600001 710000  
 WP ABA03041\_07 700001 810000  
 WP ABA03041\_08 800001 910000  
 WP ABA03041\_09 900001 1010000  
 WP ABA03041\_10 100001 1110000  
 WP ABA03041\_11 110001 1210000  
 WP ABA03041\_12 120001 1310000  
 WP ABA03041\_13 130001 1410000  
 WP ABA03041\_14 140001 1510000  
 WP ABA03041\_15 150001 1610000  
 WP ABA03041\_16 160001 1710000  
 WP ABA03041\_17 170001 1810000  
 WP ABA03041\_18 180001 1910000  
 WP ABA03041\_19 190001 2010000  
 WP ABA03041\_20 200001 2110000  
 WP ABA03041\_21 210001 2210000  
 WP ABA03041\_22 220001 2310000  
 WP ABA03041\_23 230001 2410000  
 WP ABA03041\_24 240001 2510000  
 WP ABA03041\_25 250001 2610000  
 WP ABA03041\_26 260001 2710000  
 WP ABA03041\_27 270001 2810000  
 WP ABA03041\_28 280001 2910000  
 WP ABA03041\_29 290001 2914528

Query Match 85.6%; Score 15.4; DB 6; Length 110000;  
 Best Local Similarity 94.1%; Pred. No. 2.5e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCTCTACATATAATGCGCCG 18  
 DB 34251 TGCTTACATATAATGCGCCG 34235

RESULT 5  
 AAX86622 ID AAX86622 standard; DNA; 18 BP.  
 AC AAX86622;  
 XX DT 15-OCT-1999 (first entry)  
 XX DB Probe for wild type sonic hedgehog response element.  
 XX  
 KW Sonic hedgehog response element; transcription factor; neuronal cell;  
 KW hedgehog-mediated signaling pathway; proliferation; differentiation;  
 KW proliferative disease; basal cell carcinoma; medulloblastoma; meningioma;  
 KW bone defect; polydactyly; jaw defect; rib defect; spina bifida;  
 KW familial midline defect; cyclopia; neural tube defect; probe; ss.  
 XX OS Synthetic.  
 XX PN WO9941201-A1.  
 XX PD 19-AUG-1999.  
 XX PR 11-FEB-1999; 99W0-US003112.  
 XX PR 13-FEB-1998; 98US-00023249.  
 XX OS (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
 XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.

WP ABA03041\_00 1 110000  
 PT New transcription factors involved in hedgehog-mediated signaling, used  
 PT to identify modulators for controlling expression of target genes, e.g.  
 PT for treating cancer.  
 XX PS Example 8; Page 41; 76pp; English.

XX CC The present sequence represents a probe for a wild type sonic hedgehog  
 CC response element. The specification describes a transcription factor  
 CC which is involved in a hedgehog-mediated signaling pathway, has at least  
 CC one phosphorylation site and is dephosphorylated in response to the  
 CC pathway. This transcription factor binds to a hedgehog response element  
 CC such as sequence AAX86621. Modulation of the phosphorylation of the  
 CC transcription factor is used to control expression of target genes  
 CC involved in hedgehog-mediated signaling pathway, specifically for  
 CC regulating proliferation or differentiation of neuronal cells; treatment  
 CC of proliferative diseases (specifically basal cell carcinoma, medulloblastoma and meningioma); or to inhibit bone defects (e.g.  
 CC polydactyly, jaw or rib defects, or spina bifida). Measuring the ratio of  
 CC phosphorylated:dephosphorylated forms of transcription factor is used for  
 CC diagnosis of hedgehog-mediated signaling pathway-mediated familial  
 CC midline defects (specifically cyclopia or neural tube defects).

XX SQ Sequence 18 BP; 2 A; 5 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 82.2%; Score 14.8; DB 2; Length 18;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCTACATAATGGCCG 18  
 DB 1 GTTCTACCTGatggccg 18

RESULT 6  
 AAX94156 ID AAX94156 standard; DNA; 48 BP.  
 AC AAX94156;  
 XX DT 30-JAN-2001 (first entry)  
 XX DE Adenovirus 5 E1B promoter region vector PCR primer #1.  
 XX KW Adenovirus 5; Ad 5; transcription factor Tcf binding site; colon cancer;  
 KW metastasis; liver tumour; colorectal cancer; gene therapy; PCR primer;  
 KW ss.  
 XX OS Mastadenovirus.  
 XX PN WO200056909-A1.  
 XX PD 28-SEP-2000.  
 XX PR 24-MAR-2000; 2000WO-GB001142.  
 XX PR 24-MAR-1999; 99GB-00006815.  
 XX PR (BTG) BTG INT LTD.  
 XX PI Isgro R, Brunori M;  
 XX WPI; 2000-628270/60.

XX PT Viral DNA construct for treating neoplasms comprises tumor specific  
 PT transcription factor binding sites in place of wild type transcription  
 PT factor binding sites, operatively positioned in promoter region.

XX PS Example; Page 30; 89pp; English.

XX CC The present sequence is a PCR primer used during the construction of a  
 CC vector comprising the adenovirus 5 (Ad 5) E2 and E3 transcription sites.  
 CC This sequence was mutated so that the E1B promoter was replaced with 4  
 CC copies of the Tcf transcription factor binding site. It is an example of

CC the viral sequences of the invention, which are directed at tumour cells. CC They are able to replicate so that they do not need to be given in large CC quantities or inserted directly into the tumour, and those containing the CC Tcf binding site are particularly useful in treating colon cancer. In CC addition, metastases of the cancer, such as those found in the liver and CC colorectal cancers can be treated using gene therapy in a similar way. XX

RESULT: 8  
ADR60599  
ID ADR60599 standard; cDNA; 342 BP.  
XX  
AC ADR60599;  
VV

| Qy       | 1  | GTTCATCATATGCCCG | 18 | Pred. No. 2, 1e+02;   | Length 48;                   |
|----------|----|------------------|----|---|------------------------------|
| Db       | 29 | GTCCTATATATGCCCG | 46 | Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | Best Local Similarity 88.9%; |
| RESULT 7 |    |                  |    |   |                              |
| AAA91133 |    |                  |    |   |                              |

DB Cotton cDNA sequence, SEQ ID 1380.  
 XX  
 KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;  
 KW drought tolerance; plant disease resistance; galactomannan; lignin;  
 KW plant growth regulator; heat tolerance; herbicide tolerance;  
 KW homologous recombination; extreme osmotic condition tolerance;  
 KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
 KW stress resistance.  
 XX  
 XX

XX  
AC  
AA941133;  
XX  
DT  
30-JAN-2001 (first entry)  
XX  
DE  
XX  
Adenovirus mutated E1B promoter.

XX  
PF  
XX  
PP  
07-MAY-2007 00000 00000 00000

KW  
XX  
OS  
XX  
PN  
XX  
PPD  
28-SBR-2000

Kovalic BK, Zhou Y, Cao Y, et al. *PLoS One* 2013; 8(1): e54833. doi:10.1371/journal.pone.0054833

new recombinant nucleic acid molecules and polypeptides from *Gossypium hirsutum*, useful for producing plants with improved biological characteristics (e.g. improved plant cold or drought tolerance).

Viral DNA construct for treating neoplasms comprises tumor specific transcription factor binding sites in place of wild type transcription factor binding sites, operatively positioned in promoter region.

the invention relates to a recombinant polynucleotide comprising any of the 5878 Cott plant cDNA sequences mentioned in the specification. Also a recombinant polypeptide comprising any of the 5878 amino acid sequences mentioned in the specification and producing a plant having an improved property. Producing a plant having an improved property comprises transforming a plant with a recombinant construct comprising a promoter region functional in a plant cell operably joined to a polynucleotide comprising a coding sequence for a polypeptide associated with the Cott plant.

The present sequence comprises the adenovirus E1B promoter. It has been mutated to form one of the preferred sequences of the invention, which are directed at tumour cells. These include sequences where the E2 promoter has been replaced with 4 copies of the Tcf transcription factor binding site. They are able to replicate so that they do not need to be given in large quantities or inserted directly into the tumour, and those containing the Tcf binding site are particularly useful in treating colon cancer. In addition, metastases of the cancer, such as those found in the liver and colorectal cancers can be treated using gene therapy in a similar way.

the invention relates to a recombinant polynucleotide comprising any of the 58798 Cotton plant cDNA sequences mentioned in the specification. Also a recombinant polypeptide comprising any of the 58798 amino acid sequences mentioned in the specification and producing a plant having an improved property. Producing a plant having an improved property comprises transforming a plant with a recombinant construct comprising a promoter region functional in a plant cell operably joined to a polynucleotide comprising a coding sequence for a polypeptide associated with the property, and growing the transformed plant. The polypeptide is useful for improving plant cold tolerance, manipulating growth rate in plant cells by modification of the cell cycle pathway, improving plant drought tolerance, providing increased resistance to plant disease, producing galactomannan (or lignin or plant growth regulators), improving plant heat tolerance, improving plant tolerance to herbicides, increasing the rate of homologous recombination in plants, improving plant tolerance to extreme osmotic conditions or to pathogens or pests, improving yield by modification of photosynthesis, modifying seed oil or protein yield and/or oil content, improving yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, or improving yield by providing improved plant growth and development under at least one stress condition. The polynucleotide and/or polypeptide may be provided in a recombinant construct comprising a promoter region functional in a plant cell operably joined to a polynucleotide comprising a coding sequence for a polypeptide associated with the property, and growing the transformed plant. The polypeptide is useful for improving plant cold tolerance, manipulating growth rate in plant cells by modification of the cell cycle pathway, improving plant drought tolerance, providing increased resistance to plant disease, producing galactomannan (or lignin or plant growth regulators), improving plant heat tolerance, improving plant tolerance to herbicides, increasing the rate of homologous recombination in plants, improving plant tolerance to extreme osmotic conditions or to pathogens or pests, improving yield by modification of photosynthesis, modifying seed oil or protein yield and/or oil content, improving yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, or improving yield by providing improved plant growth and development under at least one stress condition. The

Query Match      Score 14.8; DB 3; Length 180;  
 Best Local Similarity 89.9%; No. 2.5e+02;  
 Matches 16; Conservative 0; Mismatches 2;

patent did not form part of the invention. Note: The sequence data for this



|    |  |             |  |    |
|----|--|-------------|--|----|
| AC | XX   | XX          | XX   | XX |
| XX | DT   | 29-JUL-2004 | (first entry)  |    |
| XX | XX   | XX          | XX   | XX |
| DB | Pseudomonas aeruginosa polynucleotide #14361.  |             |  |    |
| XX | Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;   |             |  |    |
| KW | antibacterial.   |             |  |    |
| XX | Pseudomonas aeruginosa.  |             |  |    |
| OS | US6551795-B1.  |             |  |    |
| XX | US6551795-B1.  |             |  |    |
| PN |  |             |  |    |
| XX | 22-APR-2003.   |             |  |    |
| PD |  |             |  |    |
| XX | 18-FEB-1999; 99US-00252991.  |             |  |    |
| PF |  |             |  |    |
| XX | 18-FEB-1998; 98US-0074788P.  |             |  |    |
| PR |  |             |  |    |
| XX | 27-JUL-1998; 98US-0094190P.  |             |  |    |
| PA |  |             |  |    |
| XX | (GENO-) GENOME THERAPEUTICS CORP.  |             |  |    |
| PT |  |             |  |    |
| XX | Rubenfield MJ, Nolling J, Deloughery C, Bush D;  |             |  |    |
| DR |  |             |  |    |
| XX | WPI; 2003-615309/58.   |             |  |    |
| DR |  |             |  |    |
| XX | P-PSDB; AB082186.  |             |  |    |
| PS |  |             |  |    |
| XX | The invention relates to Pseudomonas aeruginosa polypeptides and the useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.   |             |  |    |
| XX | Disclosure; SEQ ID NO 14361; 455pp; English.   |             |  |    |
| CC |  |             |  |    |
| CC | The invention relates to Pseudomonas aeruginosa polypeptides and the useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective anti-bacterial targets, as targets for anti-bacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa nucleic acid, as components of Pseudomonas species using biochip technology. Sequences ABD01397-ABD17967 represent P. aeruginosa Polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html |             |  |    |
| CC |  |             |  |    |
| CC | Sequence 1566 BP; 283 A; 523 C; 513 G; 247 T; 0 U; 0 Other;  |             |  |    |
| CC | Query Match 82.2%; Score 14.8; DB 11; Length 1566;   |             |  |    |
| CC | Best Local Similarity 88.9%; Pred. No. 3.2e+02;  |             |  |    |
| CC | Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  |             |  |    |
| CC | QY 1 GTCCTACATATGCCCCG 18  |             |  |    |
| CC | Db 112 GCTCTATATATGCCCCG 129   |             |  |    |
| SQ |  |             |  |    |
| XX | RESULT 12  |             |  |    |
| XX | ABD15782   |             |  |    |
| XX | ABD15782 standard; DNA; 1716 BP.   |             |  |    |
| AC |  |             |  |    |
| AC | ABD15782;  |             |  |    |
| DT | 29-JUL-2004 (first entry)  |             |  |    |
| XX | Pseudomonas aeruginosa polynucleotide #14386.  |             |  |    |
| KW | Bacterial infection; gene; ds; Pseudomonas aeruginosa infection; antibacterial.  |             |  |    |
| XX |  |             |  |    |
| XX | XX   | OS          | Pseudomonas aeruginosa.  |    |
| XX | XX   | XX          | US6551795-B1.  |    |
| XX | XX   | XX          | 18-FEB-1999; 99US-00252991.  |    |
| XX | XX   | XX          | 18-FEB-1998; 98US-0074788P.  |    |
| XX | XX   | XX          | 27-JUL-1998; 98US-0094190P.  |    |
| XX | XX   | XX          | (GENO-) GENOME THERAPEUTICS CORP.  |    |
| XX | XX   | XX          | Rubenfield MJ, Nolling J, Deloughery C, Bush D;  |    |
| XX | XX   | XX          | WPI; 2003-615309/58.   |    |
| XX | XX   | XX          | P-PSDB; AB082186.  |    |
| XX | XX   | XX          | The invention relates to Pseudomonas aeruginosa polypeptides and the useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection. |    |
| XX | XX   | XX          | Disclosure; SEQ ID NO 14386; 455pp; English.   |    |
| CC |  |             |  |    |
| CC | The invention relates to Pseudomonas aeruginosa polypeptides and the useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective anti-bacterial targets, as targets for anti-bacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa nucleic acid, as components of Pseudomonas species using biochip technology. Sequences ABD01397-ABD17967 represent P. aeruginosa Polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html |             |  |    |
| CC |  |             |  |    |
| CC | Sequence 1716 BP; 269 A; 499 C; 627 G; 321 T; 0 U; 0 Other;  |             |  |    |
| CC | Query Match 82.2%; Score 14.8; DB 11; Length 1716;   |             |  |    |
| CC | Best Local Similarity 88.9%; Pred. No. 3.3e+02;  |             |  |    |
| CC | Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  |             |  |    |
| CC | QY 1 GTCCTACATATGCCCCG 18  |             |  |    |
| CC | Db 1692 GCTCTATATATGCCCCG 1709   |             |  |    |
| SQ |  |             |  |    |
| XX | RESULT 13  |             |  |    |
| XX | AD50197/c  |             |  |    |
| XX | AD50197 standard; cDNA; 3057 BP.   |             |  |    |
| AC |  |             |  |    |
| AC | AD50197;   |             |  |    |
| DT | 02-DEC-2004 (first entry)  |             |  |    |
| XX |  |             |  |    |
| DB | Bacterial polynucleotide #4940.  |             |  |    |
| XX |  |             |  |    |
| KW | Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polynucleotide; gene; ss.   |             |  |    |
| XX |  |             |  |    |
| OS | Bacteria.  |             |  |    |
| YY |  |             |  |    |



DT 16-DEC-2004 (first entry)  
 XX  
 DE Human autoimmune disease-related SNP context sequence - SEQ ID 4077.  
 XX  
 KW single nucleotide polymorphism detection; SNP detection;  
 KW rheumatoid arthritis; type 1 diabetes; multiple sclerosis;  
 KW systemic lupus erythematosus; inflammatory bowel disease; psoriasis;  
 KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;  
 KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;  
 XX KW primary systemic vasculitis; ds.  
 OS Homo sapiens.  
 XX  
 PN WO2004083403-A2.  
 XX  
 PD 30-SEP-2004.  
 XX  
 PR 18-MAR-2004; 2004WO-US008461.  
 XX  
 PR 18-MAR-2003; 2003US-045544P.  
 XX 25-APR-2003; 2003US-0465241P.  
 PA (APPL-) APPLERA CORP.  
 XX  
 PT Cargill M, Begovich AB, Alexander HC;  
 XX DR WPI, 2004-728480/71.  
 XX  
 PT New isolated nucleic acid molecule comprises at least 8 contiguous nucleotides where one of the nucleotides is a single nucleotide polymorphism (SNP), useful for diagnosing or treating autoimmune diseases, e.g. rheumatoid arthritis.  
 XX  
 PS Claim 16; SEQ ID NO 4077; 123Pp; English.  
 XX  
 CC The invention comprises amino acid and coding sequences containing genetic polymorphisms associated with an altered risk of developing an autoimmune disease (e.g. rheumatoid arthritis). The invention further comprises a method of identifying an individual that has an altered risk of developing an autoimmune disease, comprising detecting a single nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA and protein sequences of the invention are useful for diagnosing and treating autoimmune diseases, such as: rheumatoid arthritis, type 1 diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease, myocarditis, Sjogren's disease or primary systemic vasculitis. The present DNA sequence represents a human autoimmune disease-related genomic-based SNP context sequence of the invention. NOTE: The present sequence is not shown in the specification, but has been retrieved from the WIPO website.  
 XX  
 SQ Sequence 201 BP; 52 A; 37 C; 46 G; 65 T; 0 U; 1 Other;  
 Best Local Similarity 80.0%; Score 14.4; DB 13; Length 201;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 TCTTACATATGGCC 17  
 Db 125 TCTTACATATGGCC 110

Copyright (c) 1993 - 2005 Compugen Ltd.

Gencore version 5.1.6

Om nucleic - nucleic search, using sw model  
 Run on: March 2, 2005, 02:07:02 ; Search time 128 Seconds  
 (without alignments)  
 230.101 Million cell updates/sec

title: US-10-677-982-1  
 Perfect score: 18

Sequence: GTTCTACATAATGCGCCG 18

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued Patents NA: \*

1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq: \*  
 2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq: \*  
 3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq: \*  
 4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq: \*  
 5: /cgn2\_6/ptodata/1/ina/PCU-COMB.seq: \*  
 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query | Length | DB | ID                   | Description        |
|------------|-------|-------|--------|----|----------------------|--------------------|
| c 1        | 18    | 100.0 | 18     | 3  | US-09-023-249-1      | Sequence 1, Appli  |
| c 2        | 18    | 100.0 | 18     | 3  | US-09-023-249-2      | Sequence 2, Appli  |
| c 3        | 18    | 100.0 | 18     | 4  | US-09-934-035-1      | Sequence 1, Appli  |
| c 4        | 18    | 100.0 | 18     | 4  | US-09-934-035-2      | Sequence 2, Appli  |
| c 5        | 18    | 82.2  | 18     | 3  | US-09-023-249-3      | Sequence 3, Appli  |
| c 6        | 14.8  | 82.2  | 18     | 4  | US-09-934-035-3      | Sequence 3, Appli  |
| c 7        | 14.8  | 82.2  | 18     | 4  | US-09-916-510A-30    | Sequence 4, Appli  |
| c 8        | 14.8  | 82.2  | 180    | 4  | US-09-916-510A-7     | Sequence 5, Appli  |
| c 9        | 14.8  | 82.2  | 1302   | 4  | US-09-250-991A-1431  | Sequence 6, Appli  |
| c 10       | 14.8  | 82.2  | 1479   | 4  | US-09-252-991A-14286 | Sequence 7, Appli  |
| c 11       | 14.8  | 82.2  | 1566   | 4  | US-09-252-991A-14361 | Sequence 8, Appli  |
| c 12       | 14.8  | 82.2  | 1716   | 4  | US-09-252-991A-14386 | Sequence 9, Appli  |
| c 13       | 14.4  | 80.0  | 92387  | 4  | US-09-949-016-14563  | Sequence 10, Appli |
| c 14       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14568  | Sequence 11, Appli |
| c 15       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14569  | Sequence 12, Appli |
| c 16       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14570  | Sequence 13, Appli |
| c 17       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14571  | Sequence 14, Appli |
| c 18       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14572  | Sequence 15, Appli |
| c 19       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14573  | Sequence 16, Appli |
| c 20       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14574  | Sequence 17, Appli |
| c 21       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14575  | Sequence 18, Appli |
| c 22       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14576  | Sequence 19, Appli |
| c 23       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14577  | Sequence 20, Appli |
| c 24       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14578  | Sequence 21, Appli |
| c 25       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14579  | Sequence 22, Appli |
| c 26       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14580  | Sequence 23, Appli |
| c 27       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14581  | Sequence 24, Appli |
| c 28       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14582  | Sequence 25, Appli |
| c 29       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14583  | Sequence 26, Appli |
| c 30       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14584  | Sequence 27, Appli |
| c 31       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14585  | Sequence 28, Appli |
| c 32       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14586  | Sequence 29, Appli |
| c 33       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14587  | Sequence 30, Appli |
| c 34       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14588  | Sequence 31, Appli |
| c 35       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14589  | Sequence 32, Appli |
| c 36       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14590  | Sequence 33, Appli |
| c 37       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14591  | Sequence 34, Appli |
| c 38       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14592  | Sequence 35, Appli |
| c 39       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14593  | Sequence 36, Appli |
| c 40       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14594  | Sequence 37, Appli |
| c 41       | 14.4  | 80.0  | 151295 | 4  | US-09-949-016-14595  | Sequence 38, Appli |
| c 42       | 13.8  | 76.7  | 314    | 4  | US-09-949-016-14596  | Sequence 39, Appli |
| c 43       | 13.8  | 76.7  | 551    | 4  | US-09-212-991A-12375 | Sequence 40, Appli |
| c 44       | 13.8  | 76.7  | 876    | 4  | US-09-107-532A-551   | Sequence 41, Appli |
| c 45       | 13.8  | 76.7  | 987    | 4  | US-09-583-110-1890   | Sequence 42, Appli |

#### ALIGNMENTS

RESULT 1  
 US-09-023-249-1  
 Sequence 1, Application US/09023249A  
 Patent No. 6277566  
 GENERAL INFORMATION:  
 APPLICANT: Beacky, Philip A.,  
 Trai, Ming-Jer,  
 Krishnan, Venkatesh,  
 Chien, Chien-Huan  
 TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSE: Gray Cary Ware & Freidenrich LLP  
 STREET: 4365 Executive Drive, Suite 1600  
 CITY: San Diego  
 STATE: CA  
 COUNTRY: U.S.A.  
 ZIP: 92121  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: Windows95  
 SOFTWARE: Patentin Release #11.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US-09-023, 249A  
 FILING DATE: 13-Feb-1998  
 CLASSIFICATION: Application  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 09/023, 249  
 FILING DATE: 13-Feb-98  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Hale, Lisa A.  
 REGISTRATION NUMBER: 38, 347  
 REFERENCE/DOCKET NUMBER: JHU1510-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 858/677-1456  
 TELEFAX: 858/677-1465  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 18 BASE PAIRS  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 FEATURE:  
 NAME/KEY:  
 LOCATION:  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-023-249-1

Query Match 100.0%; Score 18; DB 3; Length 18; ; Patent No. 6733971  
 Best Local Similarity 100.0%; Pred. No. 0.5; ; GENERAL INFORMATION:  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0; ; APPLICANT: Beachy, Philip A.,  
 Qy 1 GTCCTACATATGCGCG 18 Tsai, Sophia Y.,  
 Db 1 GTCCTACATATGCGCG 18 Tsai, Ming-Jer,  
 Krishnan, Venkatesh,  
 Chen, Chien-Huan,  
 TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY  
 GENERAL INFORMATION:  
 NUMBER OF SEQUENCES: 8  
 APPLICANT: Beachy, Philip A.,  
 ADDRESSER: Gray Cary Ware & Freidenrich LLP  
 STREET: 4365 Executive Drive, Suite 1600  
 CITY: San Diego  
 STATE: CA  
 ZIP: 92121  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: Windows95  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/023,249A  
 CLASSIFICATION: Application  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 09/023,249  
 FILING DATE: 11-Feb-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile, Lisa A.  
 REGISTRATION NUMBER: 38-347  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 858/677-1456  
 FAX: 858/677-1465  
 INFORMATION FOR SEQ ID NO: 1:  
 REFERENCE/DOCKET NUMBER: JHU1510-1  
 LENGTH: 18 BASE Pairs  
 SEQUENCE CHARACTERISTICS:  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 FEATURE:  
 NAME/KEY:  
 LOCATION:  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-934-035-1  
 Query Match 100.0%; Score 18; DB 4; Length 18; ; Patent No. 6733971  
 Best Local Similarity 100.0%; Pred. No. 0.5; ; GENERAL INFORMATION:  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0; ; APPLICANT: Beachy, Philip A.,  
 Qy 1 GTCCTACATATGCGCG 18 Tsai, Sophia Y.,  
 Db 1 GTCCTACATATGCGCG 18 Tsai, Ming-Jer,  
 Krishnan, Venkatesh,  
 Chen, Chien-Huan,  
 TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSER: Gray Cary Ware & Freidenrich LLP  
 STREET: 4365 Executive Drive, Suite 1600  
 CITY: San Diego

RESULT 3  
 US-09-934-035-1  
 Sequence 1, Application US/09934035

RESULT 2  
 US-09-023-249-2/c  
 Sequence 2, Application US/09023249A  
 Patent No. 627756  
 GENERAL INFORMATION:  
 NUMBER OF SEQUENCES: 8  
 TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY  
 CORRESPONDENCE ADDRESS:  
 ADDRESSER: Gray Cary Ware & Freidenrich LLP  
 STREET: 4365 Executive Drive, Suite 1600  
 CITY: San Diego  
 STATE: CA  
 ZIP: 92121  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: Windows95  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/934 035  
 CLASSIFICATION: Application  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 09/023,249A  
 FILING DATE: 11-Feb-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile, Lisa A.  
 REGISTRATION NUMBER: 38-347  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 858/677-1456  
 FAX: 858/677-1465  
 INFORMATION FOR SEQ ID NO: 1:  
 REFERENCE/DOCKET NUMBER: JHU1510-1  
 LENGTH: 18 BASE Pairs  
 SEQUENCE CHARACTERISTICS:  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 FEATURE:  
 NAME/KEY:  
 LOCATION:  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-934-035-1  
 Query Match 100.0%; Score 18; DB 4; Length 18; ; Patent No. 6733971  
 Best Local Similarity 100.0%; Pred. No. 0.5; ; GENERAL INFORMATION:  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0; ; APPLICANT: Beachy, Philip A.,  
 Qy 1 GTCCTACATATGCGCG 18 Tsai, Sophia Y.,  
 Db 1 GTCCTACATATGCGCG 18 Tsai, Ming-Jer,  
 Krishnan, Venkatesh,  
 Chen, Chien-Huan,  
 TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSER: Gray Cary Ware & Freidenrich LLP  
 STREET: 4365 Executive Drive, Suite 1600  
 CITY: San Diego

STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92121

COMPUTER READABLE FORM:  
OPERATING SYSTEM: Windows95  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/934,035  
FILING DATE: 21-Aug-2001  
CLASSIFICATION: Application

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/023,24949  
FILING DATE: 1998-02-13  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: JHU1510-1

REFERENCE/DOCKET NUMBER: JHU1510-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 858/677-1456  
TELEFAX: 858/677-1465

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 BASE PAIRS  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
FEATURE:  
NAME/KEY:  
LOCATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-934-035-2

Query Match 100.0%; Score 18; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTTCTACATAATGGCCG 18  
||| ||||| ||||| |||||  
Db 1 GTTCTACATAATGGCCG 18  
||| ||||| ||||| |||||

RESULT 5  
US-09-023-249-3  
SEQUENCE 3, Application US/09023249A  
GENERAL INFORMATION:  
SEQUENCE 3, Application US/09023249A  
PATENT NO. 6277566

GENERAL INFORMATION:  
APPLICANT: Beachy, Philip A.,  
Trai, Sophia Y.,  
Trai, Ming-Jer,  
Krishnan, Venkatesh,  
Chen, Chien-Huan

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:  
APPLICANT: Beachy, Philip A.,  
Trai, Sophia Y.,  
Trai, Ming-Jer,  
Krishnan, Venkatesh,  
Chen, Chien-Huan

TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Gray Cary Ware & Freidenrich LLP  
STREET: 4365 Executive Drive, Suite 1600  
CITY: San Diego  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92121

COMPUTER READABLE FORM:  
OPERATING SYSTEM: Windows95  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/934,035  
FILING DATE: 21-Aug-2001  
CLASSIFICATION: Application

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/023,24949  
FILING DATE: 1998-02-13  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: JHU1510-1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 858/677-1456  
TELEFAX: 858/677-1465

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 BASE PAIRS  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: DNA

PRIOR APPLICATION DATA:

US-09-934-035-3

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Query Match 82.2%; Score 14.8; DB 4; Length 18;

Best Local Similarity 88.9%; Pred. No. 38; Mismatches

Matches 16; Conservative 0; Indels 0; Gaps 0;

Qy 1 GTTCTACATAATGCGCG 18

Db 1 GTTCTACAGTATGCGCG 18

RESULT 7  
US-09-916-510A-30

; Sequence 30, Application US/09916510A

; Patent No. 6544507

; GENERAL INFORMATION:

; APPLICANT: IGOO, RICHARD D.

; APPLICANT: BRONRI, MICHELE A.

; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRAL AGENTS

; FILE REFERENCE: 604-596

; CURRENT APPLICATION NUMBER: US/09/916, 510A

; PRIORITY FILING DATE: 2001-07-30

; CURRENT FILING DATE: 1999-03-24

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn Ver. 2.1

; LENGTH: 48

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Primer

; US-09-916-510A-30

Query Match 82.2%; Score 14.8; DB 4; Length 48;  
Best Local Similarity 88.9%; Pred. No. 46; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTTCTACATAATGCGCG 18

Db 29 GTCTCTATATAATGCGCG 46

RESULT 8  
US-09-916-510A-7

; Sequence 7, Application US/09916510A

; Patent No. 6544507

; GENERAL INFORMATION:

; APPLICANT: IGOO, RICHARD D.

; APPLICANT: BRONRI, MICHELE A.

; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRAL AGENTS

; FILE REFERENCE: 604-596

; CURRENT APPLICATION NUMBER: US/09/916, 510A

; PRIORITY FILING DATE: 2001-07-30

; CURRENT FILING DATE: 1999-03-24

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO: 30

; LENGTH: 180

; TYPE: DNA

; ORGANISM: Adenovirus VR5

; US-09-916-510A-7

Query Match 82.2%; Score 14.8; DB 4; Length 180;  
Best Local Similarity 88.9%; Pred. No. 58; Mismatches

Matches 16; Conservative 0; Indels 0; Gaps 0;

Qy 1 GTTCTACATAATGCGCG 18

Db 111 GTCTCTATATAATGCGCG 128

RESULT 9  
US-09-252-991A-14317/C

; Sequence 14317, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 101196-136

; CURRENT APPLICATION NUMBER: US/09/252, 991A

; PRIORITY APPLICATION NUMBER: 1998-02-18

; PRIORITY FILING DATE: 1998-02-18

; PRIORITY APPLICATION NUMBER: US 60/094, 190

; CURRENT FILING DATE: 1999-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO: 14286

; LENGTH: 1479

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-14286

Query Match 82.2%; Score 14.8; DB 4; Length 1479;  
Best Local Similarity 88.9%; Pred. No. 85; Mismatches

Matches 16; Conservative 0; Indels 0; Gaps 0;

Qy 1 GTTCTACATAATGCGCG 18

Db 158 GTCTCTATATAATGCGCG 141

RESULT 11  
US-09-252-991A-14361

; Sequence 14361, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196-136

; CURRENT APPLICATION NUMBER: US/09/252, 991A

```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14361
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-14361

Query Match 82.2%; Score 14.8; DB 4; Length 1566;
Best Local Similarity 80.9%; Pred. No. 86;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTTCTACATAATGGCCG 18
Db 112 GCCTATATATATGGCCG 129

RESULT 12
US-09-252-991A-14386
; Sequence 14386, Application US/09252991A
; Patent No. 6551795
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14386
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-14386

Query Match 82.2%; Score 14.8; DB 4; Length 1716;
Best Local Similarity 88.9%; Pred. No. 87;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTTCTACATAATGGCCG 18
Db 1692 GCTCTATATATGGCCG 1709

RESULT 13
US-09-949-016-14563
; Sequence 14563, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14563
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE: misc_feature
; NAME/KEY: misc_feature
; LOCATION: (1)..(151295)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14563

Query Match 80.0%; Score 14.4; DB 4; Length 151295;
Best Local Similarity 93.8%; Pred. No. 3.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTCTACATAATGGCCG 17
Db 73070 TTCTACATAATGGCC 73085

RESULT 14
US-09-949-016-14568
; Sequence 14568, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14568
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE: misc_feature
; NAME/KEY: misc_feature
; LOCATION: (1)..(151295)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14568

Query Match 80.0%; Score 14.4; DB 4; Length 151295;
Best Local Similarity 93.8%; Pred. No. 3.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTCTACATAATGGCCG 17
Db 73070 TTCTACATAATGGCC 73085

RESULT 15
US-09-949-016-14569
; Sequence 14569, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PULSEQ for Windows Version 4.0
; SEQ ID NO 14563
; LENGTH: 92387

Query Match 80.0%; Score 14.4; DB 4; Length 92387;
Best Local Similarity 93.8%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTCTACATAATGGCC 17
Db 73070 TTCTACATAATGGCC 73085

```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 14569
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(151295)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14569

Query Match 80.0%; Score 14.4; DB 4; Length 151295;
Best Local Similarity 93.8%; Pred. No. 3.3e+02; 1;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 TCTCTACATATAGGCC 17
Db 73070 TTTAGATATAGGCC 73085
```

Search completed: March 2, 2005, 04:45:46  
Job time : 137 sec<sub>8</sub>

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using BW model

Run on: March 2, 2005, 02:13:11 ; Search time 4752 Seconds  
(without alignments)  
22.445 Million cell updates/sec

Title: US-10-677-982-1

Perfect score: 18  
Sequence: 1 GTTCTAACATAATGCAGCG 18

Scoring table: IDENTITY\_NUC  
gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0 200000000

Maximum DB seq length: 0 200000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/us07\_pubcomb.seq:\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_pubcomb.seq:\*

4: /cgn2\_6/ptodata/2/pubpna/us07\_pubcomb.seq:\*

5: /cgn2\_6/ptodata/2/pubpna/us07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/2/pubpna/us08\_PUBCOMB.seq:\*

8: /cgn2\_6/ptodata/2/pubpna/us09\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/2/pubpna/us09c\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/2/pubpna/us09\_NEW\_PUB.seq:\*

11: /cgn2\_6/ptodata/2/pubpna/us10\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/us10\_PUBCOMB.seq:\*

13: /cgn2\_6/ptodata/2/pubpna/us10b\_PUBCOMB.seq:\*

14: /cgn2\_6/ptodata/2/pubpna/us10c\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/2/pubpna/us10d\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/2/pubpna/us10e\_PUBCOMB.seq:\*

17: /cgn2\_6/ptodata/2/pubpna/us10f\_PUBCOMB.seq:\*

18: /cgn2\_6/ptodata/2/pubpna/us10\_NEW\_PUB.seq:\*

19: /cgn2\_6/ptodata/2/pubpna/us11\_NEW\_PUB.seq:\*

20: /cgn2\_6/ptodata/2/pubpna/us11\_NEW\_PUB.seq:\*

21: /cgn2\_6/ptodata/2/pubpna/us60\_PUBCOMB.seq:\*

22: /cgn2\_6/ptodata/2/pubpna/us60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match Length | DB ID                       | Description        | ALIGNMENTS         |
|------------|-------|--------------------|-----------------------------|--------------------|--------------------|
| 1          | 18    | 100.0              | 18 9 US-07-934-035-1        | Sequence 1, Appli  | Sequence 86671, A  |
| c          | 18    | 100.0              | 18 9 US-07-934-035-2        | Sequence 2, Appli  | Sequence 3, Appli  |
| c          | 18    | 100.0              | 18 17 US-10-677-982-1       | Sequence 1, Appli  | Sequence 30, Appli |
| c          | 18    | 100.0              | 18 17 US-10-677-982-2       | Sequence 2, Appli  | Sequence 30, Appli |
| c          | 18    | 91.1               | 231 18 US-10-47-96-12834    | Sequence 12834, A  | Sequence 7, Appli  |
| c          | 16.4  | 91.1               | 18 US-10-425-115-164709     | Sequence 12849, A  | Sequence 126750, A |
| c          | 16.4  | 89.7               | 18 US-10-423-115-42326      | Sequence 42326, A  | Sequence 1380, AP  |
| c          | 8     | 85.6               | 379 18 US-10-425-115-128320 | Sequence 128320, A | Sequence 4225, A   |
| c          | 9     | 85.6               | 533 17 US-10-424-599-4803   | Sequence 4803, AP  | Sequence 74924, A  |
| c          | 15.4  | 85.6               | 918 18 US-10-437-963-10575  | Sequence 10575, AP | Sequence 112951, A |
| 11         | 15.4  | 85.6               | 957 18 US-10-767-701-9687   | Sequence 9687, AP  | Sequence 61255, A  |

RESULT 1  
US-09-934-035-1  
; Sequence 1, Application US/09934035  
; Patent No. US20020102646A1  
GENERAL INFORMATION:  
APPLICANT: Brachy, Philip A.,  
Teai, Sophia Y.,  
Tsai, Ming-Wei,  
Krishnan, Venkatesh,  
Chen, Chien-Huan.  
TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Gray Cary Ware & Freidenrich LLP  
STREET: 4365 Executive Drive, Suite 1600  
CITY: San Diego  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows95  
CURRENT SOFTWARE: Patentin Release #1.0, Version #1.30  
APPLICATION DATA:  
APPLICATION NUMBER: US/09/934,035  
FILING DATE: 21-Aug-2001  
CLASSIFICATION: Application  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 09/023,24949  
FILING DATE: 1998-02-13  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: JHUI1510-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 858/677-1456  
 TELEFAX: 858/677-1465  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 18 BASE PAIRS  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 FEATURE:  
 NAME/KEY:  
 LOCATION:  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-934-035-1

RESULT 2  
 US-09-934-035-2/c

Query Match 100.0%; Score 18; DB 9; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 5,5; Mismatches 0; Indels 0; Gaps 0;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTTCTACATATGGCGCG 18  
 Db 1 GTTCTACATATGGCGCG 18

RESULT 3  
 US-10-677-982-1

Sequence 1, Application US/10677982  
 Publication No. US2004008203A1  
 GENERAL INFORMATION:  
 APPLICANT: Beachy, Philip A.,  
 Tsai, Sophia Y.,  
 Tsai, Ming-Jer,  
 Krishnan, Venkatesh,  
 Chen, Chien-Huan  
 TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Gray Cary Ware & Freidenrich LLP  
 STREET: 4305 Executive Drive, Suite 1600  
 CITY: San Diego  
 STATE: CA  
 COUNTRY: U.S.A.  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: Windows95  
 CURRENT APPLICATION DATA:  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/334,035  
 FILING DATE: 21-Aug-2001  
 CLASSIFICATION: Application  
 PRIORITY INFORMATION:  
 APPLICATION NUMBER: 09/023,24949  
 FILING DATE: 1998-02-13  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile, Lisa A.  
 REGISTRATION NUMBER: 38,347  
 REFERENCE NUMBER: 38,347  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 858/677-1456  
 TELEFAX: 858/677-1465  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 18 BASE PAIRS  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 FEATURE:  
 NAME/KEY:  
 LOCATION:  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-677-982-1

Query Match 100.0%; Score 18; DB 17; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 5,5; Mismatches 0; Indels 0; Gaps 0;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTTCTACATATGGCGCG 18  
 Db 1 GTTCTACATATGGCGCG 18



```

; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO: 42326
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE: OTHER INFORMATION: Clone ID: MRT4577_138602C.1
; US-10-425-115-42326

Query Match 91.1%; Score 16.4; DB 18; Length 879;
Best Local Similarity 94.4%; Pred. No. 71; Mismatches 1; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GTTCTACATAATGGCCG 18
Db 583 GTTCATACATGCGCCG 566

RESULT 8
US-10-425-115-128320
; Sequence 128320, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO: 128320
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE: OTHER INFORMATION: Clone ID: MRT4577_4849C.1
; US-10-425-115-128320

Query Match 85.6%; Score 15.4; DB 18; Length 379;
Best Local Similarity 94.1%; Pred. No. 2.3e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 TTCTACATAGCGCCG 18
Db 65 TTCTACATAGCGCCG 81

RESULT 9
US-10-424-599-4803/c
; Sequence 4803, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO: 4803
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE: OTHER INFORMATION: Clone ID: PAT_MRT3847_104341C.1
; US-10-424-599-4803

Query Match 85.6%; Score 15.4; DB 18; Length 533;
Best Local Similarity 94.1%; Pred. No. 2.4e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GTTCTACATAATGGCC 17
Db 269 GTCTCATATAATGCTCC 253

RESULT 10
US-10-437-963-10575
; Sequence 10575, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204366
; SEQ ID NO: 10575
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE: OTHER INFORMATION: Clone ID: PAT_MRT4530_16883C.1
; US-10-437-963-10575

Query Match 85.6%; Score 15.4; DB 18; Length 918;
Best Local Similarity 94.1%; Pred. No. 2.4e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 TTCTACATAGCGCCG 18
Db 332 TTCTACATAGCGCCG 348

RESULT 11
US-10-757-701-9687
; Sequence 9687, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO: 9687
; LENGTH: 957
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE: OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS29420_1
; US-10-767-701-9687

Query Match 85.6%; Score 15.4; DB 18; Length 957;
Best Local Similarity 94.1%; Pred. No. 2.6e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Db 357 TCTACATCGCGCCG 373

RESULT 12

; Publication No. US20040214272A1

; Sequence 8671, Application US/10425115

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovacic, David K.

; APPLICANT: Zhou, Yihua

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21 (5322)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 363326

; SEQ ID NO: 8671

; LENGTH: 1074

; ORGANISM: Zea mays

; FEATURE: OTHER INFORMATION: CLONE ID: MRT4577\_179051C.1

US-10-425-115-8671

Query Match 85.6%; Score 15.4; DB 18; Length 1074; Best Local Similarity 94.1%; Pred. No. 2.6e+02; Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTCTACATATCGCGC 18

Db 393 TTCTACATCGCGCCG 409

RESULT 13

US-09-934-035-3

; Sequence 3, Application US/09934035

; Pattern No. US20020103646A1

; GENERAL INFORMATION:

; APPLICANT: Beachy, Philip A., Tsai, Sophia Y., Tsai, Ming-Jer, Krishnan, Venkatesh, Chan, Chien-Huan

; TITLE OF INVENTION: MODULATION OF HEDGEHOG-MEDIATED SIGNALING PATHWAY

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Gray Cary Ware & Freidenrich LLP

STREET: 4365 Executive Drive, Suite 1600

CITY: San Diego

STATE: CA

COUNTRY: U.S.A.

ZIP: 92121

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows95

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/677,982

FILING DATE: 01-Oct-2003

CLASSIFICATION: Application

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/023,249A

FILING DATE: 13-Feb-1998

APPLICATION NUMBER: 09/023,249

FILING DATE: 13-Feb-98

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: JHUI510-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 858/677-1456

TELEFAX: 858/677-1465

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 BASE PAIRS

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: Linear

MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-677-982-3

Query Match 82.2%; Score 14.8; DB 17; Length 18; Best Local Similarity 88.9%; Pred. No. 3.4e+02; Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCTACATAATGCCCG 18  
 Db 1 GTTCTACGTCATGCCCG 18

RESULT 15

US-09-916-510A-30

; Sequence 30, Application US/09916510A

; Patent No. US 20020163349A1

; GENERAL INFORMATION:

; APPLICANT: IGGO, RICHARD D.

; APPLICANT: BRUNOTTI, MICHELE A.

; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRAL AGENTS

; FILE REFERENCE: 604-596

; CURRENT APPLICATION NUMBER: US/09/916,510A

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: GB 9906815.7

; PRIOR FILING DATE: 1999-03-24

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 30

; LENGTH: 48

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE: OTHER INFORMATION: Description of Artificial Sequence: Primer

; OTHER INFORMATION: Description of Artificial Sequence: Primer

US-09-916-510A-30

Query Match 82.2%; Score 14.8; DB 9; Length 48;  
 Best Local Similarity 88.9%; Pred. No. 3.9e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCTACATAATGCCCG 18  
 Db 29 GTCTTATATAATGCCCG 46

Search completed: March 2, 2005, 06:05:04  
 Job time : 4755 secs



/note="Organ: Roots; Vector: pBluescript II SK (+) XR;  
 Site 1: EcoRI (5', end of cDNA); Site 2: XhoI (3', end of  
 cDNA); mRNA was isolated from each tissue source  
 independently and equal quantities of mRNA from each  
 tissue were then pooled. cDNA was prepared from 5  
 micrograms of mRNA and directionally ligated into the  
 pBluescript II SK (+) XR vector using the pBluescript II  
 XR cDNA Library Construction Kit according to  
 manufacturer's instructions with modifications  
 (Stratagene). Plasmid DNA was then transformed by  
 electroporation into DH10B cells (Invitrogen) for  
 propagation. Normalization was applied according to  
 published methods [Bonaldo M.F. et al. (1996) Genome  
 Research 6 (9):791] in order to reduce the abundance of  
 highly expressed transcripts."

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 497;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GTTCTACATAATGGCCG 18  
 Db 320 GTTCTACATAATGGCCG 303

## RESULT 2

CF401440  
 Locus CF401440  
 Definition RTWML\_12\_A07\_91\_A015 Well-watered loblolly pine roots WML\_Pinus  
 Accession CF401440  
 Version CF401440.1  
 Keywords EST  
 Source Pinus taeda  
 Organism Pinus taeda

REFERENCE Pratt,L., Cordomier-Pratt,M.-M., Lorenz,W.W., Dean,J.J., and  
 Neale,D.  
 AUTHORS  
 TITLE Unpublished (2003)  
 JOURNAL Other ESTs: RTWML\_12\_A07\_b1\_A015  
 COMMENT Contact: Cordomier-Pratt MM  
 The University for Genomics and Bioinformatics  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of  
 Forestry, University of Georgia; plant material prepared at the  
 University of Florida; sequencing done in the Laboratory for  
 Genomics and Bioinformatics, University of Georgia. Sequence ends  
 have been trimmed to exclude vector and regions below Phred quality  
 16. Three-prime sequences are presented as their reverse complement  
 seq primer: JENREV (CAGGAAACGCTAAGC).  
 Location/Qualifiers  
 1. .S22  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="tree 13-271"  
 /db\_xref="taxon:3330"  
 /clone="G00014b\_B01"  
 /sex="Hermaphrodite"  
 /tissue\_type="entire strobilus"  
 /dev\_stage="Three stages of preformed male strobili at end  
 of winter dormancy were pooled: swollen fully closed buds"  
 /lab\_host="E. coli DH10B cells"  
 /clone\_1tb="G00014b\_B01"  
 /note="Organ: Expanding male strobili; Vector: pBluescript  
 II SK (+) XR; Site\_1: EcoRI; Site\_2: XhoI; cDNA was  
 prepared from 5 mg of poly A+ selected RNA and was  
 directionally ligated into the pBluescript II SK (+) XR  
 vector (Stratagene), transformed by electroporation into  
 DH10B cells (Invitrogen) for propagation"

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 530;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Roots were harvested for RNA isolation. Double-stranded  
 cDNA was cloned unidirectionally into pSML180. Inserts  
 excised with EcoRI (5' end) and xhoI (3' end)."

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 522;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GTTCTACATAATGGCCG 18  
 Db 422 GTTCTACATAATGGCCG 439

## RESULT 3

CK443140  
 Locus CK443140  
 Definition G00014b\_BR\_B01\_G0001: Male strobili developmental sequence Picea  
 Accession CK443140  
 Version CK443140.1  
 Keywords EST  
 Source Picea glauca (white spruce)  
 Organism Picea glauca  
 Spermophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 (bases 1 to 530)  
 AUTHORS Morency,M.-J. and C30)  
 TITLE Arborea EST sequencing in Picea glauca (white spruce)  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: John Mackay  
 Centre de Recherche en Biologie Forestiere  
 Universite Laval  
 Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
 Fax: 418 656 7493  
 Email: j.mackay@ravs.ulaval.ca  
 Center for Computational Genomics and Bioinformatics (CCGB),  
 University of Minnesota, MN id Identifier: MN519424 Clone ID:  
 G00014b\_B01 Clones available through: John Mackay, Ph. D.  
 Professeur adjoint - Assistant professor E-MAIL:  
 j.mackay@ravs.ulaval.ca Centre de Recherche en Biologie Forestiere  
 (Forest Biology Research Center) Universite Laval Quebec, Quebec  
 CANADA G1K 7P4  
 Plate: 4b row: 01 column: B  
 Seq primer: M13 Reverse Primer.  
 Location/Qualifiers  
 1. .S30  
 /organism="Picea glauca"  
 /mol\_type="mRNA"  
 /strain="tree 13-271"  
 /db\_xref="taxon:3330"  
 /clone="G00014b\_B01"  
 /sex="Hermaphrodite"  
 /tissue\_type="entire strobilus"  
 /dev\_stage="Three stages of preformed male strobili at end  
 of winter dormancy were pooled: swollen fully closed buds"  
 /lab\_host="E. coli DH10B cells"  
 /clone\_1tb="G00014b\_B01"  
 /note="Organ: Expanding male strobili; Vector: pBluescript  
 II SK (+) XR; Site\_1: EcoRI; Site\_2: XhoI; cDNA was  
 prepared from 5 mg of poly A+ selected RNA and was  
 directionally ligated into the pBluescript II SK (+) XR  
 vector (Stratagene), transformed by electroporation into  
 DH10B cells (Invitrogen) for propagation"

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 530;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/note="Vector: pSML180; Site\_1: EcoRI; Site\_2: XhoI; The  
 library was prepared from polyA+ RNA from loblolly pine  
 (Pinus taeda) roots watered to pot capacity every other  
 day. Pre-dawn water potential remained -0.3 MPa +/-0.1.

FEATURES  
 source  
 1. .S22  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCLONES"  
 /db\_xref="taxon:3332"  
 /clone="RTWML\_12\_A07\_A015"  
 /lab\_host="DH5<sup>TM</sup>-T1 phage-resistant E. coli"  
 /clone\_1tb="Well-watered loblolly pine roots WML"  
 /note="Vector: pSML180; Site\_1: EcoRI; Site\_2: XhoI; The  
 library was prepared from polyA+ RNA from loblolly pine  
 (Pinus taeda) roots watered to pot capacity every other  
 day. Pre-dawn water potential remained -0.3 MPa +/-0.1.

|   |                     |  |   |   |
|---|---------------------|--|---|---|
| Qy  | 1                   | GTCTACATAATGCCCG 18                                  | VERSION   | CK443141.1  |
| Db  | 317                 | GTCTACATAATGCCCG 334                                 | EST   |   |
| RESULT  | 4                   |  | KEYWORDS  | Picea glauca; white spruce)   |
| LOCUS   | CF401389            | 573 bp mRNA linear EST 29-AUG-2003                   | SOURCE  |   |
| DEFINITION  | RTW1_12_A07_b1_A015 | Well-watered loblolly pine roots WW1_Pinus           | ORGANISM  | Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.                                     |
| ACCESSION   | CF401389            | taeda_cDNA clone RTW1_12_A07_A015 3', mRNA sequence. | REFERENCE   | 1 (bases 1 to 608)  |
| VERSION   | CF401389.1          | EST  | AUTHORS   | Morency,M.-J., Cooke,J., Pavly,N., Parsons,J., Paule,C., Seguin,A., Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J., Siddiqui,A., Holt,R., Marra,M. and Mackay,J. |
| COMMENT   |                     | Unpublished (2003)                                   | JOURNAL   | Arboresc EST sequencing in Picea glauca (white spruce)  |
| KEYWORDS  |                     |  | COMMENT   | Unpublished (2004)  |
| SOURCE  |                     |  | CONTACT   | Contact: John Mackay/   |
| ORGANISM  |                     |  | Centre de Recherche en Biologie Forestiere  |   |
| Pinus taeda (loblolly pine)   |                     |  | Universite Laval  |   |
| Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.   |                     |  | Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  |   |
| REFERENCE   |                     |  | Fax: 418 656 7493   |   |
| AUTHORS   |                     |  | Email: jmackay@rbsv.ulaval.ca   |   |
| Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Dean,J.,   |                     |  | Center for Computational Genomics and Bioinformatics (CCGB), University of Minnesota, MN 1d Identifier: MN515425 Clone ID: G0014b B01   |   |
| Gebremedhin,M., Dervinis,C., Martin,T., White,T., Davis,J. and Neale,D.   |                     |  | Clones available, through: John Mackay, Ph. D.  |   |
| TITLE   |                     |  | Professeur adjoint-Assistant professor EMAIL: Jmackay@rbsv.ulaval.ca Centre de Recherche en Biologie Forestiere (Forrest Biology Research Center) Universite Laval Quebec, Quebec   |   |
| JOURNAL   |                     |  | Plate: 4b row: 01 column: B   |   |
| COMMENT   |                     |  | Seq primer: PolyTplus Primer.   |   |
| KEYWORDS  |                     |  | Location/Qualifiers   |   |
| SOURCE  |                     |  | I. . 608  |   |
| ORGANISM  |                     |  | /organism="Picea glauca"  |   |
| RNA, prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. seq primer: M13-21 (TGTAAAACGACGGCCAGT) POLYA>No. |                     |  | /mol_type="mRNA"  |   |
| FEATURES  |                     |  | /strain="Tree 13-271"   |   |
| SOURCE  |                     |  | /db_xref="taxon:3330"   |   |
| ORGANISM  |                     |  | /clone="GQ0014b B01"  |   |
| PRIMER  |                     |  | /sex="Hermaphrodite"  |   |
| FEATURES  |                     |  | /tissue_type="Entire strobilus"   |   |
| SOURCE  |                     |  | /dev_stage="Three stages of preformed male strobili at end of winter dormancy were pooled: swollen fully closed buds, partly open buds and fully open buds"   |   |
| ORGANISM  |                     |  | /lab_host="E. coli DH10B cells"   |   |
| PRIMER  |                     |  | /clone_lis="GQ001: Male strobili developmental sequence"  |   |
| FEATURES  |                     |  | /notes="Organ: Expanding male strobili; Vector: pBluecript II SK (+)XR, Site 1: Eco-RI, Site 2: Xba-I, cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluecript II SK (+)XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitro) for propagation" |   |
| SOURCE  |                     |  | ORIGIN  |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |
| FEATURES  |                     |  |   |   |
| SOURCE  |                     |  |   |   |
| ORGANISM  |                     |  |   |   |
| PRIMER  |                     |  |   |   |





The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: [impratt@uga.edu](mailto:impratt@uga.edu)  
 RNA prepared and library constructed by W. Walter Lorenz, School of  
 Forestry, University of Georgia; Plant material prepared at the  
 University of Florida; sequencing done in the laboratory for  
 Genomics and Bioinformatics, University of Georgia. Sequence ends  
 have been trimmed to exclude vector and regions below Phred quality  
 16. Three-prime sequence ends are presented as their reverse complement  
 and have been trimmed to exclude polyA.  
 Seq primer: JENREV (CAGCTAACAGCTATGACC).

## FEATURES

## Source

Location/Qualifiers  
 1. . .377  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCLONES"  
 /db\_xref="raxon:3352"  
 /clone="RIW1\_3\_G12\_A015"  
 /lab\_host="DH10-T1 phage-resistant E. coli"  
 /clone\_lib="Well-washed loblolly pine roots WW1"  
 /notes="Vector: pSL180; Site 1: EcoRI; Site 2: XbaI; The  
 library was prepared from PolyA+ RNA from loblolly pine  
 (Pinus taeda) roots watered to pot capacity every other  
 day. Pre-dawn water potential remained -0.3 MPa +/-0.1.  
 Roots were harvested for RNA isolation. Double-stranded  
 cDNA was cloned unidirectionally into pSL180. Inserts  
 excised with EcoRI (5' end) and XbaI (3' end)."

## ORIGIN

## Source

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

## 0;

## Gaps

## 0;

## Length

## 737;

## Query

## Match

## Best Local Similarity

## Matches

## Score

## DB

## Pred.

## No.

## Mismatches

## 0;

## Indels

/organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /tissue\_type="leaf and stem, including leaf base"  
 /dev\_stage="2 week old seedling (3 leaves)"  
 /lab\_host="X11-Blue"  
 /clone\_lib="947 - 2 week shoot from Barkan lab"  
 /note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-);  
 site\_1: EcoRI; Site\_2: XbaI; Directionally cloned using  
 Stratagene's UniZap XR cDNA cloning kit with the 5' end  
 at the EcoRI site. The library represents 8 x 10<sup>5</sup>  
 independent recombinant phage. The plants were greenhouse  
 grown."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 4; Length 303;  
 Best Local Similarity 94.4%; Pred. No. 2.e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GTTCTACATTAATGCGCCG 18  
 Db 44 GTTCTACATCATGCGCCG 61

## RESULT 13

LOCUS CA150538  
 DEFINITION SCBFRR2019C07.g RZ2 Saccharum officinarum cDNA clone SCBFRR2019C07  
 ACCESSION CA150538  
 VERSION CA150538.1  
 KEYWORDS EST  
 SOURCE Saccharum officinarum  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 REFERENCE 1 (bases 1 to 515)  
 AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.I. and Arruda,P.  
 TITLE The libraries that made SUCEST  
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 COMMENT Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parruda@camp.br  
 Clone distribution: clone distribution information can be found  
 through the Brazilian Clone Collection Center (BCCC) at  
 http://www.bcccenter.fcav.unesp.br  
 Plate: 019 Row: C column: 07  
 Seq Primer: T7 Promoter Primer.

## COMMENT

JOURNAL  
 An EST database from well-watered loblolly pine (*Pinus taeda*) roots  
 unpublished (2003)  
 Other ESTs: RTW1\_3\_G12.g1\_A015  
 Contact: Cordeanier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7221, USA  
 Tel: 706 542 1860  
 Fax: 706 503 0200  
 Email: mmp@uga.edu  
 RNA prepared and library constructed by W. Walter Lorenz, School of  
 Forestry, University of Georgia; plant material prepared at the  
 University of Florida; sequencing done in the Laboratory for  
 Genomics and Bioinformatics, University of Georgia. Sequence ends  
 have been trimmed to exclude vector and regions below Phred quality  
 16. Three-prime sequences are presented as their reverse complement  
 and have been trimmed to exclude polyA.  
 Seq Primer: M13-21 (TGTAAAACGACGCCAGT)  
 POLYA=Yes.

FEATURES  
 source  
 location/Qualifiers

1. .515

/organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 /clone="SCBFRR2019C07"  
 /lab\_host="DH10B"  
 /clone\_lib="RZ2"  
 /note="Organ: Shoot-root transition zone from young plants  
 Site\_2: Ncol; An unidirectional cDNA library generated  
 from [Shoot-root transition zone from young plants (small  
 insert library)]. cDNA was prepared from polyA+ mRNA  
 using SuperScript Plasmid System Kit (Invitrogen). The  
 double-strand cDNAs were fractionated in a sepharose  
 CL-2B 40cm-columns and fragments sizing between 0.8 and  
 1.5 Kb were directionally cloned into the vector. details  
 of each source of RNA and library construction can be  
 obtained at <http://sucest.lad.ic.unicamp.br/public>"

FEATURES  
 source

1. .478

/organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCLONEs"  
 /db\_xref="taxon:3352"  
 /clone="RTW1\_3\_G12\_A015"  
 /lab\_host="DH10B-11 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots WW1"  
 /note="Vector: pSL180; Site\_1: EcoRI; Site\_2: XbaI; The  
 library was prepared from polyA+ RNA from loblolly pine  
 (Pinus taeda) roots watered to pot capacity every other  
 day. Pre-dawn water potential remained -0.3 mpa +/-0.1.

Roots were harvested for RNA isolation. Double-stranded  
 cDNA was cloned unidirectionally into pSL180. Inserts  
 excised with EcoRI (5' end) and XbaI (3' end)."

## ORIGIN

Query Match 91.1%; Score 16.4; DB 7; Length 478;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GTTCTACTAATGCGCCG 18  
 Db 18 GTTCTACTAATGCGCCG 35

## RESULT 14

LOCUS CA150538  
 DEFINITION SCBFRR2019C07.g RZ2 Saccharum officinarum cDNA clone SCBFRR2019C07  
 REFERENCE 1 (bases 1 to 515)  
 AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.I. and Arruda,P.  
 TITLE The libraries that made SUCEST  
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 COMMENT Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parruda@camp.br  
 Clone distribution: clone distribution information can be found  
 through the Brazilian Clone Collection Center (BCCC) at  
 http://www.bcccenter.fcav.unesp.br  
 Plate: 019 Row: C column: 07  
 Seq Primer: T7 Promoter Primer.

FEATURES  
 source

1. .515

/organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 /clone="SCBFRR2019C07"  
 /lab\_host="DH10B"  
 /clone\_lib="RZ2"  
 /note="Organ: Shoot-root transition zone from young plants  
 Site\_2: Ncol; An unidirectional cDNA library generated  
 from [Shoot-root transition zone from young plants (small  
 insert library)]. cDNA was prepared from polyA+ mRNA  
 using SuperScript Plasmid System Kit (Invitrogen). The  
 double-strand cDNAs were fractionated in a sepharose  
 CL-2B 40cm-columns and fragments sizing between 0.8 and  
 1.5 Kb were directionally cloned into the vector. details  
 of each source of RNA and library construction can be  
 obtained at <http://sucest.lad.ic.unicamp.br/public>"

## ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 515;  
 Best Local Similarity 94.4%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## QY

1 GTTCTACTAATGCGCCG 18

Db 208 GTTCTACTAATGCGCCG 225

Search completed: March 2, 2005, 03:05:19  
Job time : 3130 secs

RESULT 15  
BB12144  
LOCUS BE121144 524 bp mRNA linear EST 13-JUN-2000  
DEFINITION UI-R-CA0-bau-e-11-0-UI-81 UI-R-CA0 Rattus norvegicus cDNA clone  
ACCESSION UI-R-CA0-bau-e-11-0-UI 3', mRNA sequence.  
VERSION BE121144.1 GI:8513249  
KEYWORDS BE121144.1  
EST  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 524)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97054477  
PUBMED 8880548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 8565  
Email: bento-soares@uiowa.edu  
The sequence contained an oligo-dT track that was present in the  
strand cDNA and therefore this may represent a bona fide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to identify it as a clone from the  
normalized hippocampus library cDNA library preparation: M.B. the  
Soares Lab Clone distribution: clones will be available through  
Research Genetics (www.resgen.com) The following repetitive  
elements were found in this cDNA sequence: 180-208,  
>(CAAA)n#Simple repeat  
>Seq\_Primer: M13 Forward  
Seq\_Primer: M13 Forward  
POLYA=Yes

FEATURES Source

FEATURES Location/Qualifiers

1. .524 /organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="Taxon:10116"  
/clone="UI-R-CA0-bau-e-11-0-UI"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="UI-R-CA0"  
/note="Vector: PT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CA0  
library is a subtracted library derived from the following  
tissues: thalamus, cerebellum, hypothalamus, medulla,  
pons, midbrain, cerebral cortex, corpus striatum, testis,  
and hippocampus. For a detailed description of the  
library from which this clone was derived, please visit  
our web site at ratest.eugr.uiova.edu. The subtraction  
has been previously described in (Bonaldo, Lennon and  
Soares, Genome Research 6:791-806, 1996)  
TAG\_LIB=UI-R-CA0  
TAG\_SEQ=GATTC  
TAG\_SHQ=GATTC"

ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 524;  
Best Local Similarity 94.4%; Prid. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GATCTACATAATGGCCG 18  
Db 380 GCTCTACATATGGCCG 397